PATENT

UTSD:553

APPLICATION FOR UNITED STATES LETTERS PATENT

for

ABSORPTIVE HYPERCALCIURIA LOCUS ON HUMAN CHROMOSOME 1

by

Berenice Y. Reed-Gitomer

and

Charles Y.C. Pak

BACKGROUND OF THE INVENTION

This application is a continuation of U.S. Provisional Application, S.N. 60/090,348, filed June 23, 1998. The government owns rights in the present invention pursuant to grant numbers PO1-DK20543 and MO1-RR00633 from the National Institute of Health.

1. Field of the Invention

The present invention relates generally to the field of hypercalciuria. More particularly, it concerns the determination and identification of a genetic basis for absorptive hypercalciuria and osteoporosis with hypercalciuria. This determination allows the development of diagnostics and therapeutics.

2. Description of Related Art

Nephrolithiasis is a common debilitating clinical disorder associated with an estimated lifetime risk of stone formation of 20% for males and of 5% for females in the western population. In the United States, the annual incidence of nephrolithiasis is 7 to 21 per 10,000, with up to 10% of patients undergoing active stone passage requiring hospitalization to control complications.

Absorptive hypercalciuria (AH) causes stone formation in about 50% of the reported cases. AH is invariably associated with intestinal hyperabsorption of calcium in the presence of normal serum calcium concentration and a normal or suppressed level of parathyroid hormone. Osteoporosis or bone loss, particularly of trabecular bone (Barkin *et al.,* 1985), is a frequent complication. The mechanism by which hypercalciuria leads to osteoporosis is not fully understood.

Both clinical and experimental data indicate that AH is heterogeneous in origin. Serum calcitriol concentration is high in some, but not all, patients with AH (Kaplan *et al.,* 1977, Broadus *et al.,* 1984, Breslau *et al.,* 1992). The reduced calcitriol synthesis

with ketoconazole restores normal intestinal calcium absorption in some patients, but not in all of them (Breslau *et al.,* 1992). While spinal bone density is often low, some patients enjoy normal density. Also some patients with AH exhibit exaggerated renal synthesis of $1,25(OH)_2D$ (Insogna *et al.,* 1985).

One mechanism that has been proposed for the basis of AH involves an increase in the number of vitamin D receptors in the intestine. Li *et al.* demonstrated such an increase in the intestine of a normocalcemic, normal calcitriolemic rat model for AH (Li *et al.,* 1993). An elevated level of vitamin D receptors was also observed in the activated blood lymphocytes of some AH patients who had normal levels of circulating $1,25(OH)_2D$ (Zerwekh *et al.,* 1993). While evidence of genetic linkage between AH and the vitamin D receptor or the 1, α-hydroxylase gene loci has been pursued, none has been produced to suggest any linkage. Other reports implicate vitamin D-independent factors. A ketoconazole study showed some patients to be ketoconazole-resistant because their intestinal hyperabsorption of calcium and hypercalciuria were unaffected by treatment (Breslau *et al.,* 1992). Bianchi *et al.* suggested that an activation of the plasma membrane of Ca/ATPase may be etiologically important in AH, based on the finding of accelerated activity of this enzyme in red blood cells (Bianchi *et al.,* 1988).

While environmental-nutritional factors have been implicated in the pathogenesis of AH (Hess *et al.,* 1993), strong evidence suggests involvement of a genetic process in AH; a familial pattern is present in 45% or reported stone cases, and an autosomal dominant inheritance pattern has been disclosed (Coe *et al.,* 1979, Pak *et al.,* 1981). Stone formation may be influenced by multiple risk factors, both environmental and intrinsic. However, the intrinsic factors, that is, a molecular and genetic basis of AH, have not yet been characterized. The evaluation of large stone-forming kindreds by Coe *et al.* (1979) and by the group (Pak *et al.*, 1981) indicated that AH was inherited in an autosomal dominant manner. However, no molecular genetic basis for the intestinal hyperabsorption of calcium in AH has been identified. It has been speculated that AH could result from stimulation of renal $1,25(OH)_2D$ synthesis, (Insogna *et al.*, 1985;

Broadus *et al.*, 1984) increased vitamin D receptor sensitivity, (Breslau *et al.*, 1992; Li *et al.*, 1993; Zerwekh *et al.*, 1993) or activation of the plasma membrane Ca/ATPase (Bianchi *et al.*, 1988). The prior studies failed to show an abnormal vitamin D receptor genotype (Zerwekh *et al.*, 1995) or a positive linkage between AH and gene loci expected to be involved vitamin D metabolism (Reed *et al.*, 1996). In Dent's disease and related conditions that have a clinical presentation that includes hypercalciuria and nephrolithiasis, a mutation in the chloride transporter gene, CLCN5, has been reported (Lloyd *et al.*, 1996). However, AH, unlike Dent's Disease, does not have an X-linked mode of inheritance.

There has been a clear need to identify, if any, a genetic basis for AH. Such information would yield a better understanding of the condition, providing important implications in the diagnosis and therapy of AH. Furthermore, identification of a genetic basis for AH may furnish a definitive diagnostic test to identify at risk, but asymptomatic, individuals. Detection of such individuals then allows for dietary and therapeutic intervention to prevent the onset of stone disease or osteoporosis.

AH is frequently accompanied by bone loss or "osteoporosis". Osteoporosis is defined as a group of disorders that is characterized by aberrant bone remodeling; the net rate of bone resorption is greater, rather than in dynamic equilibrium with, the rate of bone formation. The condition can occur as either a primary disorder or as a disorder associated with a various disease, such as hypercalciuria. Examples of osteoporosis with hypercalciuria include ideopathic osteoporosis with hypercalciuria and postmenopausal osteoporosis with hypercalciuria. Ideopathic ostoeporosis is often times seen in young women or men demonstrating increased calcium absorption for unknown reasons. Postmenopausal osteoporosis is seen in postmenopausal women and is associated with decreased estrogen levels and increased calcium absorption. L2-L4 bone density was 10% below normal levels overall and had declined by more than 25% in approximately one-fourth of patients who had AH (Pietchmann *et al.*, 1992). Histomorphometric studies confirmed an abnormal bone picture, characterized by an increased osteoclastic resorption

surface (Bordier *et al.,* 1977), decreased osteoblastic activity (Malluche *et al.,* 1980), or both (Steiniche *et al.,* 1989) in idiopathic hypercalciuria. In the animal model of AH, bone calcium loss was associated with an increase in Vitamin D receptor (Krieger *et al.,* 1996). However, the exact role of bone in the pathogenesis of AH remains unclear. Some have implicated cytokine involvement in the etiology of bone loss associated with AH (Pacifici *et al.,* 1990 ;Weisinger *et al.,* 1996). Increased production of IL-1 by monocytes of hypercalciuric patients has been observed (Pacifici *et al.,* 1990), and a recent study reported that hypercalciuric patients with stones had increased levels of basal secretion of IL-1α by circulating monocytes and enhanced levels of TNFα and IL-6 production by activated monocytes (Weisinger *et al.,* 1996). It was suggested that IL-1 could stimulate prostaglandin production, which would account for previous reports of high PGE_2 in hypercalciuric stone-formers (Buck *et al.,* 1981). Weisinger subsequently implied that prostaglandin-dependent synthesis of calcitriol could cause hyperabsorption of calcium and hypercalciuria, but no concrete evidence of this has been produced.

An understanding of a molecular/genetic basis of AH would facilitate the development of new therapeutic strategies for the treatment of AH. Current diagnosis is based on stone risk profile, markers for bone turnover, and bone densitometry and also involves blood tests and urinalysis. The results of many of these tests are influenced significantly by diet and thus require patient compliance with defined diet. A straightforward genetic test would eliminate the complications of extended testing and increase the certainty of the diagnosis. Unfortunately such a test is not presently available.

SUMMARY OF THE INVENTION

The present invention relates to the inventors' discovery that there exists an area on human chromosome 1 that is genetically linked to absorptive hypercalciuria (AH), and thus to some forms of osteoporosis as well. The invention further relates to the

development of a familial screening method based on the identification of this region on human chromosome 1. The invention also contemplates further refining the locus to identify a gene involved in AH and to use this information for familial and nonfamilial detection and therapeutic intervention of AH and osteoporosis with hypercalciuria.

The discoveries of this invention eliminate the complications associated with screening methods currently available. Extensive blood tests and urinalysis that require compliance with a defined diet are no longer necessary to implement the present invention, which involves a simple, straightforward genetic test that can be implemented in diagnosing AH and osteoporosis with hypercalciuria.

Described in this invention is a method for screening for an increased risk of hypercalciuria by obtaining a sample nucleic acid from a subject; and analyzing the sample nucleic acid to detect the presence or absence of a genetic mutation in genomic region associated with an increased risk of developing hypercalciuria. The hypercalciuria is further defined as absorptive hypercalciuria or as osteoporosis with hypercalciuria. In certain embodiments, the osteoporosis with hypercalciuria is further defined as ideopathic osteoporosis with hypercalciuria or postmenopausal osteoporosis with hypercalciuria.

In one embodiment, this invention contemplates the identification of a region associated with an increased risk of AH on human chromosome 1. In further embodiments, the region is defined as containing 1q23 and 1q24. Further refinement of this region identifies a loci between markers D1S2681 and D1S2815. This region is further defined as having a lod score of greater than 3.0 but less than 30.0. Moreover, markers D1S2681 and D1S2815 further define a 4.3 megabase region when analyzed against two kindreds. The genomic region associated with an increased risk of AH may have a sequence contained in at least one genetic sequence selected from the group consisting of the the genetic sequences set forth in GenBank Accession # Z97876, GenBank Accession # Z99943, and GenBank Accession # AL031733.

The genomic region associated with an increased risk of AH contains a large number of genes encoding a large number of proteins. Within this large genomic regions may be families of genes, related to each other structurally or functionally. Isoforms of related proteins or groupings of subunits of a larger multimeric protein structure are often times found localized to the same genomic region. Therefore, the genetic lesions actually associated with an increased risk of AH may localize to more than one gene in this area. It is expected that there are several unique mutations associated with an increased risk of AH in different individuals.

In one aspect of the invention, a putative AH-genetically associated gene and the encoded AH-genetically associated protein has been identified. This putative gene is shown as SEQ ID NO:1 and the encoded protein is shown as SEQ ID NO:2. Methodologies for identification and characterization of this putative gene and encoded protein, mapping of specific mutations associated with this putative gene, diagnostic uses of this putative gene and encoded proteins, as well as therapies and screens for modulators of this putative gene and encoded protein are also described.

Identification of a region and a marker associated with an increased risk of absorptive hypercalciuria leads to another embodiment of the present invention for the detection of family members who may be individuals at risk for AH. In certain embodiments, linkage analysis is the screening tool using markers linked to AH. In other embodiments, the detection involves screening for a mutation in the region associated with AH. These embodiments contemplate the use of PCR, hybridization techniques using a probe complementary to a portion of the region associated with an increased risk of AH, and other techniques involving the detection or characterization of nucleic acids.

The nucleic acid to be analyzed can be either RNA or DNA. The nucleic acid can be analyzed from a blood sample, from a urine sample, or from any other tissue containing DNA or RNA corresponding to the AH loci.

The present invention discloses all loci directly related to the genetic basis for AH. This includes the nucleic acid sequences of all genes and open reading frames associated with the genetic basis for AH.

The present invention provides DNA segments, vectors and the like comprising at least a first isolated gene, DNA segment or coding sequence region that encodes a protein, polypeptide, domain, peptide or any fusion protein thereof associated with the genetic basis of AH, and particularly, that encode a human protein, domain, fragment or derivative associated with the genetic basis of AH. Moreover, all nucleic acids isolated from the AH loci are also considered in the present invention, including primers, probes, oligonucleotides, any moiety from 15 base pairs (bps) or greater, or any other distinct and discrete segment of nucleic acid that is substantially similar to the sequence of the region containing the AH locus.

As used herein in the context of the instant compositions, the AH locus providing a genetic basis for AH will be understood to include wild-type, polymorphic and mutant sequences of this region. Wild-type sequences are defined as the first identified sequence, polymorphic sequences are defined as naturally occurring variants of the wild-type sequence that have no effect on the expression or function of proteins or domains thereof associated with the genetic basis of AH, and mutant sequences are defined as changes in the wild-type sequence, either naturally occurring or introduced by the hand of man, that have an effect on either the expression, stability, cellular location, post-translation modification, and/or the function of the proteins or domains thereof that are associated with the genetic basis of AH.

Thus, the invention also includes the provision of DNA segments, vectors, genes and coding sequence regions that encode proteins, polypeptides, domains, peptides or any fusion protein thereof that are associated with the genetic basis of AH (hereinafter "AH-genetically-associated" region), where the protein element comprises at least one mutation in comparison to the wild-type sequence. The mutation may be deliberately

introduced by the hand of man, for example, in order to test the function of the changed amino acid. Additionally, the mutation may be a naturally occurring polymorphic change, either isolated from normal cells or introduced by the hand of man.

A mutation may also be in a purified protein obtained directly from an aberrant cell, or may be a recombinant protein that has been changed to introduce a mutation that mirrors one identified in a patient. A mutation may result in a gene or protein related to the genetic basis of AH, or may result in increased, decreased or undetectable levels of such a gene or protein being produced. Where diagnostic or prognostic genes, proteins and antibodies that are associated with the genetic basis of AH are concerned the mutant gene, DNA segment, antibody or even peptide will preferably have specificity for the mutant sequence in preference to the wild-type sequence, allowing effective differentiation between the two, as may be used in diagnostic or prognostic tests for AH or osteoporosis with hypercalciuria.

It will be understood that while the normal, native, wild-type proteins associated with the genetic basis of AH are defined in terms of these properties and domains, the overall features will generally be the same for AH polymorphic and mutant proteins and domains as well. The polymorphic and mutant AH-genetically-associated genes and proteins can be understood with reference to the wild-type sequences and the exemplary mutants included herein.

The genes and DNA segments of the present invention preferably encode wild-type or polymorphic proteins, polypeptides, domains, peptides or fusion constructs thereof that are associated with the genetic basis of AH where the sequence includes a contiguous amino acid sequence from the region containing the AH locus or a biologically functional equivalent thereof. The present invention also provides genes and DNA segments that encode mutant proteins, polypeptides, domains, peptides or fusion constructs thereof that are related to the genetic basis of AH where the sequence includes a contiguous amino acid sequence from a region containing the AH locus, or a

biologically functional equivalent thereof. As used herein, the term "contiguous amino acid sequence" will be understood to include a contiguous amino acid sequence of at least about 4, about 6, about 9, about 10, about 12, about 15 or about 20 amino acids or any number of amino acids greater than 20.

The DNA segments and coding regions may encode wild-type, polymorphic or mutant AH-genetically-associated peptides, *e.g.*, of from about 15 to about 30 or about 50 amino acids in length or any number of amino acids greater than 20. The peptides may be lacking in any defined AH-genetically-associated protein activity, and may, for example, be used in generating antibodies or in other embodiments. The peptides or domains may also be deliberately engineered to include a mutation, *e.g.*, in order to prepare antibodies that are specific for a mutated AH-genetically-associated gene, particularly where the mutation represents one identified in a patient AH or osteoporosis with hypercalciuria.

The present invention also provides DNA segments and coding regions that may encode an AH-genetically-associated peptide from about 6 to about 30 amino acids in length, the peptide having an amino acid sequence that corresponds to a wild-type AH-genetically associated sequence of an AH-genetically-associated protein sequence region that is susceptible to mutations that are indicative of a malignant phenotype. Where diagnostic or prognostic AH-genetically-associated genes, proteins and antibodies are concerned the gene, DNA segment, antibody or even peptide will preferably allow effective differentiation between the mutant AH-genetically-associated sequence and the wild-type AH-genetically-associated sequence as may be used in diagnostic or prognostic tests for AH or osteoporosis with hypercalciuria, as described in more detail herein below.

The genes, DNA segments, vectors and coding sequence regions may also encode wild-type, polymorphic or mutant AH-genetically-associated polypeptides and peptides with certain, but necessary all, AH-genetically-associated functional properties. As such

genes and coding sequences encoding isolated wild-type, polymorphic or mutant AH-genetically-associated domains are provided.

The AH-genetically-associated domains may also be mutant domains, which include naturally occurring polymorphisms, mutations found in AH-genetically-associated proteins in patients and, also, mutations deliberately engineered into a domain to test their function in assays. The mutant domains are also useful in antibody generation and in various *in vitro* and cellular assays. Engineering increased binding to AH-genetically-associated domains is also contemplated.

DNA segments, isolated genes or coding regions may also be manipulated to encode AH-genetically-associated fusion proteins or constructs in which at least one AH-genetically-associated protein sequence is operatively attached or linked to at least one distinct, selected amino acid sequence. This includes the combination of AH-genetically-associated sequences with selected antigenic amino acid sequences; selected non-antigenic carrier amino acid sequences; selected adjuvant sequences; amino acid sequences with specific binding affinity for a selected molecule; and amino acid sequences that form an active DNA binding or transactivation domain are particularly contemplated. Certain fusion proteins may be linked together via a protease-sensitive peptide linker, allowing subsequent easy separation.

The DNA segments intended for use in expression will be operatively positioned under the control of, *i.e.*, downstream from, a promoter that directs expression of AH-genetically-associated gene or genes in a desired host cell, such as *E. coli*, or in certain preferred embodiments in a mammalian or human cell. The promoter may be a recombinant promoter or a promoter naturally associated with an AH-genetically-associated gene. Recombinant vectors thus form another aspect of the present invention.

The use of isolated AH-genetically-associated genes positioned, in reverse orientation, under the control of a promoter that directs the expression of an antisense product in a cell is also contemplated.

The nucleic acid segments provided by the invention are thus further characterized as including:

(a) a nucleic acid segment comprising a sequence region that consists of at least about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 225, about 250, about 275, about 300, about 325, about 350, about 375, about 400 , about 450, about 475, about 500, about 550, about 600, about 650, about 700, about 750, about 800, about 850, about 900, about 950, or about 1000 contiguous nucleotides that have the same sequence as, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 225, about 250, about 275, about 300, about 325, about 350, about 375, about 400 , about 450, about 475, about 500, about 550, about 600, about 650, about 700, about 750, about 800, about 850, about 900, about 950, or are complementary to, or any other

number of contiguous nucleotides of an AH-genetically-associated nucleic acid sequence; or

(b) a nucleic acid segment from about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 225, about 250, about 275, about 300, about 325, about 350, about 375, about 400 , about 450, about 475, about 500, about 550, about 600, about 650, about 700, about 750, about 800, about 850, about 900, about 950, to about 1000, or any other number greater than about 1000 to 20,000 nucleotides in length that specifically hybridizes to the nucleic acid segment of an AH-genetically-associated sequence, or the complements thereof, under standard stringency, or preferably, under high stringency hybridization conditions.

Standard and high stringency hybridization conditions are well known to those of skill in the art. An exemplary, but not limiting, standard hybridization is incubated at 42°C in 50% formamide solution containing dextran sulfate for 48 hours and subjected to a final wash in 0.5X SSC, 0.1% SDS at 65°C. In addition to hybridization to Southern or northern blots, hybridization of primers for use in PCR is another preferred method for identification of sequences contemplated for use in the present invention.

Where the "complement" of any of the above nucleic acid segments are provided, such a complement may be functionally considered as an antisense nucleic acid, which includes nucleic acid segments positioned, in reverse orientation, under the control of a promoter that directs the expression of an antisense product. Antisense products may be

used to inhibit the transcription or translation of any AH-genetically-associated genes, in *in vitro* systems in order to more precisely define the cellular consequence of inhibition, or even *in vivo* in situations where inhibition of one or more of any AH-genetically-associated genes would be believed to be result in a beneficial effect, such as an anti-AH effect.

Mutants of each of the foregoing sequences and their encoded proteins, polypeptides, and peptides are also contemplated. The mutants may be used in the detection of physiologically relevant mutations or in further testing an functional analyses.

Segments of AH-genetically-associated nucleic acid sequences, or the complements thereof, or the mutants thereof, may variously be any length, for example, about 10,11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 or 1000 or so nucleotides in length, up to and including the full length sequences, or even longer, as may be achieved by duplication of certain domains.

Any segment may be combined into a DNA segment or vector of, for example, up to about 50,000, about 40,000, about 30,000, or about 20,000 basepairs in length. Segments of up to about 20,000, 19,000, 18,000, 17,000, 16,000, 15,000, 14,000, 13,000, 12,000, 11,000, 10,000, 9,000, 8,000, 7,000, 6,000, or about 5,000 basepairs in length will generally be preferred, and segments of up to about 5,000, 4,000 and 3,000 basepairs in length are also provided.

The nucleic acids of the present invention may also be DNA segments or RNA segments. Nucleic acid detection kits are also provided.

The present invention further provides recombinant host cells comprising at least one DNA segment or vector that comprises an isolated gene that encodes an AH-

genetically-associated protein, polypeptide, domain, peptide or any fusion protein or mutant thereof. The invention also contemplates recombinant host cells comprising at least one DNA segment or vector that comprises an isolated region of DNA that affects the transcription of a gene associated with the genetic basis for AH. Prokaryotic recombinant host cells, such as *E. coli*, are provided, as are eukaryotic host cells, including kidney cells and cells involved with bone remodeling provided with the AH-genetically-associated gene constructs of the invention.

The recombinant host cells may further comprise an operative AH-genetically-associated protein or active fragment or domain thereof. Such recombinant host cells may be provided with the AH-genetically-associated sequence *in vitro*, for example, to test its interactions, or may naturally express AH-genetically-associated sequences, including cells provided with an AH-genetically-associated sequence *in vivo* and *in vitro*, either for treatment or for study.

The recombinant host cells of the present invention preferably have one or more DNA segments introduced into the cell by means of a recombinant vector, and preferably express the DNA segment to produce the encoded AH-genetically-associated protein or peptide.

Methods of using AH-genetically-associated DNA segments are provided that comprise expressing a AH-genetically-associated DNA segment in a recombinant host cell and collecting the AH-genetically-associated protein, peptide, domain or mutant expressed by said cell. As represented by the steps of:

(a) preparing a recombinant vector in which an AH-genetically-associated-encoding DNA segment is positioned under the control of a promoter;

(b) introducing said recombinant vector into a recombinant host cell;

(c) culturing the recombinant host cell under conditions effective to allow expression of an AH-genetically-associated protein, peptide, domain or mutant; and

(d) collecting said expressed AH-genetically-associated protein, peptide, domain or mutant.

Methods for detecting AH-genetically-associated genes in cells or samples are also provided and generally comprise contacting sample nucleic acids from a sample suspected of containing an AH-genetically-associated sequence with a nucleic acid segment that encodes an AH-genetically-associated protein or peptide under conditions effective to allow hybridization of substantially complementary nucleic acids, and detecting the hybridized complementary nucleic acids thus formed.

Other methods for detecting AH-genetically-associated genes in cells or samples include amplification and PCR, RNA mismatch cleavage assays and RNase protection assays. Several variations and improvements to these assays are described in the specification and are incorporated herein by reference.

The methods may be diagnostic of AH or osteoporosis with hypercalciuria by detecting AH-genetically-associated mutants as opposed to wild-type sequences. The use of both AH-genetically-associated wild-type and mutant sequences as probes or primers in such methods will naturally be included. A wild-type sequence probe or primer will be expected to bind to the native, non-mutant sequences, but not to a mutant, and *vice versa*. The use of a mutant-specific probe that corresponds to a mutant identified in a family member with AH may be preferred in screening other family members. In any event, irrespective of the AH-genetically-associated nucleic acid segment employed, these studies will still only allow hybridization of substantially complementary nucleic acids,

thus facilitating the detection only of wild-type or only mutant hybridized nucleic acid complexes.

In further embodiments, the present invention AH-genetically-associated proteins, polypeptides, domains, peptides, mutants and any fusion proteins thereof, including AH-genetically-associated compounds purified from natural sources, such as from mammalian and human cells, and AH-genetically-associated amino acid sequences prepared by recombinant means. Recombinant AH-genetically-associated proteins and peptides may be defined as being prepared by expressing an AH-genetically-associated protein or peptide in a recombinant host cell and purifying the AH-genetically-associated protein or peptide away from total recombinant host cell components.

The AH-genetically-associated protein compositions, whether natural or recombinant, will generally be obtained free from total cell components, and will comprise at least one type of isolated AH-genetically-associated protein or peptide, purified relative to the natural level in a given cell.

AH-genetically-associated fusion proteins or constructs comprising AH-genetically-associated sequences operatively attached to distinct, selected amino acid sequences, such as selected antigenic amino acid sequences, amino acid sequences with selected binding affinity, and DNA binding or transactivation amino acid sequences, are also encompassed within the invention. Fusion proteins with selectably-cleavable bonds are also provided.

The AH-genetically-associated proteinaceous compositions will include the same types of mutants as described above for the nucleic acids. The use of specific mutated AH-genetically-associated peptides to prepare mutant-specific antibodies is particularly contemplated. In terms of diagnostic AH-genetically-associated peptides and antibodies, these compositions will generally be more useful in regard to point mutants, whereas

nucleic acid probes may be more suitable for detecting deletion, duplication, translocation and insertional mutations in addition to point mutants.

The AH-genetically-associated proteins, polypeptides, domains, peptides and fusion proteins, as well as AH-genetically-associated DNA segments, vectors, isolated genes and coding sequences may also be formulated with a pharmaceutically acceptable diluent or vehicle to form an AH-genetically-associated-pharmaceutical composition in accordance with this invention.

Further compositions of the present invention are antibodies, including monoclonal antibodies and antibody conjugates, that have immunospecificity for an AH-genetically-associated protein or peptide. The antibodies may be operatively attached to a detectable label. The antibodies and antibody conjugates may be specific for mutant AH-genetically-associated proteins or peptides and allow differential binding from wild-type AH-genetically-associated proteins. Antibody detection kits are also provided.

Certain methods for detecting AH-genetically-associated protein sequences in a sample comprise contacting a sample suspected of such sequences with a first antibody that binds to an AH-genetically-associated protein or peptide, or a mutant thereof, under conditions effective to allow the formation of immune complexes, and detecting the immune complexes thus formed. In addition to their diagnostic use, these methods are also suitable for purifying AH-genetically-associated protein sequences, identifying AH-genetically-associated protein expression, in identifying engineered mutants and in titering AH-genetically-associated proteins and/or AH-genetically-associated antibodies.

The invention further provides diagnostic methods, particularly useful in connection with AH, but also of potential usefulness in connection with osteoporosis with hypercalciuria.

Diagnostically, the present invention provides methods for identifying a patient having or at risk for AH or osteoporosis with hypercalciuria, comprising determining the type or amount of AH-genetically-associated protein present within a biological sample from the patient, wherein the presence of an AH-genetically-associated mutant or an altered amount of wild-type AH-genetically-associated protein, in comparison to a sample from a normal subject, is indicative of a patient having or at risk for AH or osteoporosis with hypercalciuria.

The "type" of AH-genetically-associated protein may be determined, allowing mutant genes and proteins to be distinguished from wild-types. The use of mutant- and wild-type-specific nucleic acid probes is particularly contemplated. In the beginning, the use of wild-type-specific nucleic acid probes will be preferred. The identification of a particularly diagnostic mutant sequence will then lead to the increased use of that mutant sequence, either in the population or in defined families. The use of mutant- and wild-type-specific antibodies is also contemplated, as may be prepared using mutant- and wild-type-specific AH-genetically-associated peptides.

Where the "amount" of AH-genetically-associated protein is determined, a differential amount of the natural AH-genetically-associated protein may be indicative of the propensity to AH or osteoporosis with hypercalciuria. Changes from the naturally observed range in the population will be easily detected and will have implications for disease risk and development.

The type or amount of AH-genetically-associated protein may be determined by means of a molecular biological assay to determine the type or amount of a nucleic acid that encodes an AH-genetically-associated protein. Such molecular biological assays will often comprise a direct or indirect step that allows a determination of the sequence of at least a portion of the AH-genetically-associated-encoding nucleic acid, which sequence can be compared to a wild-type AH-genetically-associated sequence.

It is contemplated that AH-genetically-associated sequences diagnostic or prognostic for AH or osteoporosis with hypercalciuria may comprise at least one point mutation, deletion, translocation, insertion, duplication or other aberrant change. Diagnostic RFLPs are thus also contemplated. RNase protection assays may also be employed in certain embodiments.

Diagnostic methods may be based upon the steps of:

(a) obtaining a blood or urine sample from a subject or patient;

(b) contacting sample nucleic acids from the sample with an isolated AH-genetically-associated nucleic acid segment under conditions effective to allow hybridization of substantially complementary nucleic acids; and

(c) detecting, and optionally further characterizing, the hybridized complementary nucleic acids thus formed.

The methods may involve *in situ* detection of sample nucleic acids located within the cells of the sample. The sample nucleic acids may also be separated from the cell prior to contact. The sample nucleic acids may be DNA or RNA.

The methods may involve the use of isolated AH-genetically-associated nucleic acid segments that comprises a radio, enzymatic or fluorescent detectable label, wherein the hybridized complementary nucleic acids are detected by detecting the label.

PCR® will often be preferred, as exemplified by the steps of:

(a) contacting the sample nucleic acids with a pair of nucleic acid primers that hybridize to distant sequences from a mutant, polymorphic or wild-type

AH-genetically-associated nucleic acid sequence, the primers capable of amplifying a mutant, polymorphic or wild- AH-genetically-associated nucleic acid segment when used in conjunction with a polymerase chain reaction;

(b) conducting a polymerase chain reaction to create amplification products; and

(c) detecting and characterizing the amplification products thus formed.

Diagnostic immunoassay methods are also provided, wherein the type or amount of AH-genetically-associated protein is determined by means of an immunoassay to determine the type or amount of an AH-genetically-associated protein. Such methods may comprise the steps of:

(a) obtaining a blood or urine sample from a subject or patient;

(b) contacting the sample with a first antibody that binds to an AH-genetically-associated protein or peptide, or mutant, under conditions effective to allow the formation of specific immune complexes; and

(c) detecting the specific immune complexes thus formed.

The first antibody may be linked to a detectable label, wherein the immune complexes are directly detected by detecting the presence of the label. The immune complexes may also be indirectly detected by means of a second antibody linked to a detectable label, the second antibody having binding affinity for the first antibody.

Methods utilizing the nucleic acid region that provides the genetic basis for AH are also contemplated by the present invention. Such methods include, but are not limited

to: detecting mutations in the AH nucleic acid region; characterizing any and all mutations within the region in an effort to elucidate the molecular mechanisms underlying AH and osteoporosis with hypercalciuria; identifying modulators, both in wild-type and mutant forms, that may interact with the AH nucleic acid region to alter expression of the region, to affect the integrity of the region, or to affect the stability of the region.

The invention also encompasses a method of treating individuals who have been identified as exhibiting the genetic profile of those at risk for AH or osteoporosis with hypercalciuria. In certain embodiments, the osteoporosis with hypercalciuria is further defined as ideopathic osteoporosis with hypercalciuria or postmenopausal osteoporosis with hypercalciuria.

The AH or osteoporosis with hypercalciuria treatment methods of the present invention may be combined with any standard strategy, such as dietary modification, hormone therapy, and pharmacological treatments. Examples of appropriate treatment regimes known to those of skill in the art are described in the examples. These include conservative dietary and fluid regimens to be incorporated into the daily routine of patients with kidney stones and therapeutic measures directed towards reducing urinary calcium excretion and decreasing intestinal calcium bioavalability. The treatment of asymptomatic diagnosed individuals is directed towards prevention of the first stone-forming episode. The administration of a biologically effective amount of an AH-genetically-associated protein, peptide or recombinant vector composition is also contemplated.

Furthermore, the invention considers the use of the AH nucleic acid region to encode a polypeptide, in part or whole, in order to study further the molecular basis AH and investigate diagnostic and therapeutic strategies. The identification, isolation, characterization of any modulators, either in their wild-type or mutant form, of the

polypeptide are also part of the present invention. In addition, the genetic basis for the modulators is also encompassed within the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

FIG. 1A, FIG. 1B, and FIG. 1C. Pedigrees of 3 kindreds studied. A filled symbol indicates affected individual, an open symbol represents non-evaluated individual, a "U" in a symbol represents unaffected status, a "?" in a symbol represents unknown status and an "S" below the patient identifier number indicates the presence of a kidney stone. Probands are indicated by an arrow. A slash through the symbol indicates the individual is deceased. (FIG. 1A) Pedigree for kindred AH-01. (FIG. 1B) Pedigree for kindred AH-02. (FIG. 1C) Pedigree of kindred AH-03.

FIG. 2A and FIG. 2B. Localization of the gene for AH. (FIG. 2A) Filled portion of the vertical bar indicates the interval likely to harbor the AH gene based on haplotypes. Individuals are designated as in FIG. 1. Recombinants localize the defective gene to a 4.3 cM region between D1S2681 and D1S2815, shown as the filled region of the locus bar. (FIG. 2B) Multipoint analysis: The position of marker D1S426 was arbitrarily set at 0 cM and the positions of the other loci were fixed according to composite map distance from the linkage data base. Multipoint non-parametric LOD scores on the x-axis are plotted against chromosomal-1 loci on the y-axis.

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present invention concerns the detection, diagnosis, prognosis, and treatment of absorptive hypercalciuria. The present invention describes the first genetic basis for AH. The genetic region linked to AH is disclosed. Also disclosed is a method of screening for and detecting an increased risk for hypercalciuria premised on the genetic basis for the condition. Identification of at risk individuals can result in the diagnosis and therapeutic treatment of such patients in an effort to reduce the risk of stone formation or osteoporosis.

I. Absorptive Hypercalciuria Genes and DNA Segments

Important aspects of the present invention concern isolated DNA segments and recombinant vectors encoding wild-type, polymorphic or mutant proteins whose nucleic acid sequences are genetically associated with absorptive hypercalciuria. This invention also includes the creation and use of recombinant host cells through the application of DNA technology, that express wild-type, polymorphic or mutant AH-genetically-associated proteins, using sequences located on human chromosome 1 that are genetically associated with AH.

The present invention concerns DNA segments, isolatable from mammalian and human cells, that are free from total genomic DNA and that are capable of expressing a protein or polypeptide that is encoded for by a nucleic acid sequence that is genetically associated with AH.

As used herein, the term "DNA segment" refers to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a region genetically associated with AH refers to a DNA segment that contains wild-type, polymorphic or mutant AH-genetically-associated coding sequences yet is isolated away from, or purified free from, total mammalian or human genomic DNA. Included within the term "DNA segment" are DNA segments and smaller fragments of

such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phage, viruses, and the like.

Similarly, a DNA segment comprising an isolated or purified wild-type, polymorphic or mutant AH-genetically-associated gene refers to a DNA segment including wild-type, polymorphic or mutant AH-genetically-associated protein coding sequences and, in certain aspects, wild-type, polymorphic or mutant AH-genetically associated regulatory sequences, including promoters, enhancers, and 3' regulatory sequences, isolated substantially away from other naturally occurring genes or protein encoding sequences. In this respect, the term "gene" is used for simplicity to refer to a functional protein, polypeptide or peptide encoding unit. As will be understood by those in the art, this functional term includes both genomic sequences, cDNA sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, domains, peptides, fusion proteins and mutants.

"Isolated substantially away from other coding sequences" means that the gene of interest, in this case the wild-type, polymorphic or mutant AH-genetically-associated gene forms the significant part of the coding region of the DNA segment, and that the DNA segment does not contain large portions of naturally-occurring coding DNA, such as large chromosomal fragments or other functional genes or cDNA coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

In particular embodiments, the invention concerns isolated DNA segments and recombinant vectors incorporating DNA sequences that encode a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide that includes within its amino acid sequence a contiguous amino acid sequence in accordance with the nucleic acid sequence corresponding to wild-type, polymorphic or mutant human AH-genetically-associated proteins. Moreover, in other particular embodiments, the invention concerns isolated DNA segments and recombinant vectors that encode an AH-genetically-associated

protein or peptide that includes within its amino acid sequence the substantially full length protein sequence encoded for by the nucleic acid sequence of a region genetically associated with AH.

In a preferred embodiment, a putative AH-genetically associated gene and the encoded AH-genetically associated protein has been identified. This putative gene is shown as SEQ ID NO:1 and the encoded protein is shown as SEQ ID NO:2. These sequences are given as examples of an AH-genetically associated gene and protein. Elsewhere in this application, SEQ ID NO:1 and SEQ ID NO:2 are refered to as an AH-genetically associated gene and protein, respectively.

The term "a sequence essentially as set forth in SEQ ID NO:2" means that the sequence substantially corresponds to a portion of SEQ ID NO:2 and has relatively few amino acids that are not identical to, or a biologically functional equivalent of, the amino acids of SEQ ID NO:2.

The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein. Accordingly, sequences that have between about 60% and about 65%; or more preferably, between about 66% and about 70%; or more preferably, between about 71% and about 75%; or more preferably, between about 76% and about 80%; or more preferably, between about 81% and about 85%; or more preferably, between about 86% and about 90%; or even more preferably, between about 91% and about 95%; or, between about 96% and about 99% of amino acids that are identical or functionally equivalent to the amino acids of SEQ ID NO:2 will be sequences that are "essentially as set forth in SEQ ID NO:2", provided the biological activity of the protein is maintained. Similarly, sequences that have between about 60% and about 65%; or more preferably, between about 66% and about 70%; or more preferably, between about 71% and about 75%; or more preferably, between about 76% and about 80%; or more preferably, between about 81% and about 85%; or more preferably, between about 86% and about 90%; or even more preferably, between about 91% and about 95%; or,

between about 96% and about 99% of nucleic acids that are identical or functionally equivalent to the nucleic acids of SEQ ID NO:1 will be sequences that are "essentially as set forth in SEQ ID NO:1", provided the biological activity of the sequence is maintained.

The term "functionally equivalent codon" is used herein to refer to codons that encode the same amino acid, such as the six codons for arginine or serine, and also refers to codons that encode biologically equivalent amino acids (see Table 1, below).

TABLE 1

CODON TABLE

Amino Acids			Codons
Alanine	Ala	A	GCA GCC GCG GCU
Cysteine	Cys	C	UGC UGU
Aspartic acid	Asp	D	GAC GAU
Glutamic acid	Glu	E	GAA GAG
Phenylalanine	Phe	F	UUC UUU
Glycine	Gly	G	GGA GGC GGG GGU
Histidine	His	H	CAC CAU
Isoleucine	Ile	I	AUA AUC AUU
Lysine	Lys	K	AAA AAG
Leucine	Leu	L	UUA UUG CUA CUC CUG CUU
Methionine	Met	M	AUG
Asparagine	Asn	N	AAC AAU
Proline	Pro	P	CCA CCC CCG CCU
Glutamine	Gln	Q	CAA CAG
Arginine	Arg	R	AGA AGG CGA CGC CGG CGU
Serine	Ser	S	AGC AGU UCA UCC UCG UCU
Threonine	Thr	T	ACA ACC ACG ACU
Valine	Val	V	GUA GUC GUG GUU
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAC UAU

It will also be understood that amino acid and nucleic acid sequences may include additional residues, such as additional N- or C-terminal amino acids or 5' or 3' sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological protein activity where protein expression is concerned. The addition of terminal sequences particularly applies to nucleic acid sequences that may, for example, include various non-coding sequences flanking either of the 5' or 3' portions of the coding region or may include various internal sequences, *i.e.*, introns, which are known to occur within genes.

Excepting intronic or flanking regions, and allowing for the degeneracy of the genetic code, sequences that have between about 70% and about 79%; or more preferably, between about 80% and about 89%; or even more preferably, between about 90% and about 99%; of nucleotides that are identical to the nucleotides of SEQ ID NO:1, will be sequences that are "essentially as set forth in SEQ ID NO:1"

Sequences that are essentially the same as those set forth in SEQ ID NO:1 may also be functionally defined as sequences that are capable of hybridizing to a nucleic acid segment containing the complement of SEQ ID NO:1 under relatively stringent conditions. Suitable relatively stringent hybridization conditions will be well known to those of skill in the art, as disclosed herein.

Naturally, the present invention also encompasses DNA segments that are complementary, or essentially complementary, to the sequence set forth in SEQ ID 1. Nucleic acid sequences that are "complementary" are those that are capable of base-pairing according to the standard Watson-Crick complementarity rules. As used herein, the term "complementary sequences" means nucleic acid sequences that are substantially complementary, as may be assessed by the same nucleotide comparison set forth above, or as defined as being capable of hybridizing to the nucleic acid segment of SEQ ID NO:1 under relatively stringent conditions such as those described herein.

The nucleic acid segments of the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

For example, nucleic acid fragments may be prepared that include a short contiguous stretch identical to or complementary to the nucleic acid sequence genetically associated with AH, such as about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 35, about 40, about 45, about 50, about 55, about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 225, about 250, about 275, about 300, about 325, about 350, about 375, about 400 , about 450, about 475, about 500, about 550, about 600, about 650, about 700, about 750, about 800, about 850, about 900, about 950, to about 1000, or any other number greater than about 1000 to 20,000 nucleotides in length, or about 10,000, or about 5,000 base pairs in length, with segments of about 3,000 being preferred in certain cases. DNA segments with total lengths of about 1,000, about 500, about 200, about 100 and about 50 base pairs in length (including all intermediate lengths) are also contemplated to be useful.

It will be readily understood that "intermediate lengths", in these contexts, means any length between the quoted ranges, such as 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, etc.; 21, 22, 23, etc.; 30, 31, 32, etc.; 50, 51, 52, 53, etc.; 100, 101, 102, 103, etc.; 150, 151, 152, 153, etc.; including all integers through the 200-500; 500-1,000; 1,000-2,000; 2,000-3,000; 3,000-5,000; 5,000-10,000 ranges, up to and including sequences of about 12,001, 12,002, 13,001, 13,002, 15,000, 20,000 and the like.

The various probes and primers designed around the disclosed nucleotide sequences of the present invention may be of any length. By assigning numeric values to a sequence, for example, the first residue is 1, the second residue is 2, *etc.*, an algorithm defining all primers can be proposed:

n to n + y

where n is an integer from 1 to the last number of the sequence and y is the length of the primer minus one, where n + y does not exceed the last number of the sequence. Thus, for a 10-mer, the probes correspond to bases 1 to 10, 2 to 11, 3 to 12 ... and so on. For a 15-mer, the probes correspond to bases 1 to 15, 2 to 16, 3 to 17 ... and so on. For a 20-mer, the probes correspond to bases 1 to 20, 2 to 21, 3 to 22 ... and so on.

It will also be understood that this invention is not limited to the particular nucleic acid and amino acid sequences of the region on human chromosome 1 that is genetically associated with AH. This includes the AH-genetically associated gene sequence of SEQ ID NO:1 and the AH-genetically associated protein sequence of SEQ ID NO:2. Recombinant vectors and isolated DNA segments may therefore variously include these coding regions themselves, coding regions bearing selected alterations or modifications in the basic coding region, or they may encode larger polypeptides that nevertheless include such coding regions or may encode biologically functional equivalent proteins or peptides that have variant amino acids sequences.

The DNA segments of the present invention encompass biologically functional equivalent AH-genetically-associated proteins and peptides. Such sequences may arise as a consequence of codon redundancy and functional equivalency that are known to occur naturally within nucleic acid sequences and the proteins thus encoded. Alternatively, functionally equivalent proteins or peptides may be created via the application of recombinant DNA technology, in which changes in the protein structure may be engineered, based on considerations of the properties of the amino acids being exchanged. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques, *e.g.*, to introduce improvements to the antigenicity of the protein or to test mutants in order to examine DNA binding activity at the molecular level.

One may also prepare fusion proteins and peptides, *e.g.*, where AH-genetically-associated protein coding regions are aligned within the same expression unit with other

proteins or peptides having desired functions, such as for purification or immunodetection purposes (*e.g.*, proteins that may be purified by affinity chromatography and enzyme label coding regions, respectively).

Encompassed by the invention are DNA segments encoding relatively small peptides, such as, for example, peptides of from about 15 to about 50 amino acids in length, and more preferably, of from about 15 to about 30 amino acids in length; and also larger polypeptides up to and including proteins corresponding to the full-length sequences set forth in SEQ ID NO:2.

B. Recombinant Vectors, Host Cells and Expression

Recombinant vectors form important further aspects of the present invention. The term "expression vector or construct" means any type of genetic construct containing a nucleic acid coding for a gene product in which part or all of the nucleic acid encoding sequence is capable of being transcribed. The transcript may be translated into a protein, but it need not be. Thus, in certain embodiments, expression includes both transcription of a gene and translation of a RNA into a gene product. In other embodiments, expression only includes transcription of the nucleic acid, for example, to generate antisense constructs.

Particularly useful vectors are contemplated to be those vectors in which the coding portion of the DNA segment, whether encoding a full length protein or smaller peptide, is positioned under the transcriptional control of a promoter. A "promoter" refers to a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a gene. The phrases "operatively positioned", "under control" or "under transcriptional control" means that the promoter is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the gene.

The promoter may be in the form of the promoter that is naturally associated with a wild-type, polymorphic or mutant AH-genetically-associated gene as may be obtained by isolating the 5' non-coding sequences located upstream of the coding segment or exon, for example, using recombinant cloning and/or PCR technology, in connection with the compositions disclosed herein (PCR technology is disclosed in U.S. Patent 4,683,202 and U.S. Patent 4,682,195, each incorporated herein by reference).

In other embodiments, it is contemplated that certain advantages will be gained by positioning the coding DNA segment under the control of a recombinant, or heterologous, promoter. As used herein, a recombinant or heterologous promoter is intended to refer to a promoter that is not normally associated with a wild-type, polymorphic or mutant AH-genetically-associated gene in its natural environment. Such promoters may include promoters normally associated with other genes, and/or promoters isolated from any other bacterial, viral, eukaryotic, or mammalian cell.

Similarly, the promoter or other regulatory sequences of the AH-genetically-associated region may be used to control the expression of a heterologous gene, such as a reporter gene for use in expression assays. Such genes may include genes normally associated with other bacterial, viral, eukaryotic, or mammalian cells.

Naturally, it will be important to employ a promoter that effectively directs the expression of the DNA segment in the cell type, organism, or even animal, chosen for expression or to employ a cell type, organism, or even animal that can be used with the regulatory region of a gene in the AH locus. The use of promoter and cell type combinations for protein expression is generally known to those of skill in the art of molecular biology, for example, see Sambrook *et al.* (1989), incorporated herein by reference. The promoters employed may be constitutive, or inducible, and can be used under the appropriate conditions to direct high level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant proteins or peptides.

At least one module in a promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 late genes, a discrete element overlying the start site itself helps to fix the place of initiation.

Additional promoter elements regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the tk promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either co-operatively or independently to activate transcription.

The particular promoter that is employed to control the expression of a nucleic acid is not believed to be critical, so long as it is capable of expressing the nucleic acid in the targeted cell. Thus, where a human cell is targeted, it is preferable to position the nucleic acid coding region adjacent to and under the control of a promoter that is capable of being expressed in a human cell. Generally speaking, such a promoter might include either a human or viral promoter. Preferred promoters include those derived from HSV, including the HNF1α promoter. Another preferred embodiment is the tetracycline controlled promoter.

In various other embodiments, the human cytomegalovirus (CMV) immediate early gene promoter, the SV40 early promoter and the Rous sarcoma virus long terminal repeat can be used to obtain high-level expression of transgenes. The use of other viral or mammalian cellular or bacterial phage promoters which are well-known in the art to

achieve expression of a transgene is contemplated as well, provided that the levels of expression are sufficient for a given purpose. Tables 2 and 3 below list several elements/promoters which may be employed, in the context of the present invention, to regulate the expression of wild-type, polymorphic or AH-genetically-associated gene. This list is not intended to be exhaustive of all the possible elements involved in the promotion of transgene expression but, merely, to be exemplary thereof.

Enhancers were originally detected as genetic elements that increased transcription from a promoter located at a distant position on the same molecule of DNA. This ability to act over a large distance had little precedent in classic studies of prokaryotic transcriptional regulation. Subsequent work showed that regions of DNA with enhancer activity are organized much like promoters. That is, they are composed of many individual elements, each of which binds to one or more transcriptional proteins.

The basic distinction between enhancers and promoters is operational. An enhancer region as a whole must be able to stimulate transcription at a distance; this need not be true of a promoter region or its component elements. On the other hand, a promoter must have one or more elements that direct initiation of RNA synthesis at a particular site and in a particular orientation, whereas enhancers lack these specificities. Promoters and enhancers are often overlapping and contiguous, often seeming to have a very similar modular organization.

Additionally any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of a transgene. Use of a T3, T7 or SP6 cytoplasmic expression system is another possible embodiment. Eukaryotic cells can support cytoplasmic transcription from certain bacterial promoters if the appropriate bacterial polymerase is provided, either as part of the delivery complex or as an additional genetic expression construct.

TABLE 2

PROMOTER TABLE

PROMOTER
Immunoglobulin Heavy Chain
Immunoglobulin Light Chain
T-Cell Receptor
HLA DQ α and DQ β
β-Interferon
Interleukin-2
Interleukin-2 Receptor
MHC Class II 5
MHC Class II HLA-DRα
β-Actin
Muscle Creatine Kinase
Prealbumin (Transthyretin)
Elastase *I*
Metallothionein
Collagenase
Albumin Gene
α-Fetoprotein
α-Globin
β-Globin
c-fos
c-HA-ras
Insulin
Neural Cell Adhesion Molecule (NCAM)
α_1-Antitrypsin
H2B (TH2B) Histone

PROMOTER
Mouse or Type I Collagen
Glucose-Regulated Proteins (GRP94 and GRP78)
Rat Growth Hormone
Human Serum Amyloid A (SAA)
Troponin I (TN I)
Platelet-Derived Growth Factor
Duchenne Muscular Dystrophy
SV40
Polyoma
Retroviruses
Papilloma Virus
Hepatitis B Virus
Human Immunodeficiency Virus
Cytomegalovirus
Gibbon Ape Leukemia Virus

T0370

TABLE 3

ENHANCER TABLE

Element	Inducer
MT II	Phorbol Ester (TPA) Heavy metals
MMTV (mouse mammary tumor virus)	Glucocorticoids
β-Interferon	poly(rI)X poly(rc)
Adenovirus 5 E2	ElA
c-jun	Phorbol Ester (TPA), H_2O_2
Collagenase	Phorbol Ester (TPA)

Element	Inducer
Stromelysin	Phorbol Ester (TPA), IL-1
SV40	Phorbol Ester (TPA)
Murine MX Gene	Interferon, Newcastle Disease Virus
GRP78 Gene	A23187
α-2-Macroglobulin	IL-6
Vimentin	Serum
MHC Class I Gene H-2kB	Interferon
HSP70	Ela, SV40 Large T Antigen
Proliferin	Phorbol Ester-TPA
Tumor Necrosis Factor	FMA
Thyroid Stimulating Hormone α Gene	Thyroid Hormone

Turning to the expression of the wild-type, polymorphic or mutant AH-genetically-associated proteins of the present invention , once a suitable clone or clones have been obtained such as SEQ ID NO:1, whether they be cDNA based or genomic, one may proceed to prepare an expression system. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. It is believed that virtually any expression system may be employed in the expression of the proteins of the present invention.

Both cDNA and genomic sequences are suitable for eukaryotic expression, as the host cell will generally process the genomic transcripts to yield functional mRNA for translation into protein. Generally speaking, it may be more convenient to employ as the recombinant gene a cDNA version of the gene. It is believed that the use of a cDNA version will provide advantages in that the size of the gene will generally be much smaller and more readily employed to transfect the targeted cell than will a genomic gene, which will typically be up to an order of magnitude larger than the cDNA gene.

However, the inventor does not exclude the possibility of employing a genomic version of a particular gene where desired.

In expression, one will typically include a polyadenylation signal to effect proper polyadenylation of the transcript. The nature of the polyadenylation signal is not believed to be crucial to the successful practice of the invention, and any such sequence may be employed. Preferred embodiments include the SV40 polyadenylation signal and the bovine growth hormone polyadenylation signal, convenient and known to function well in various target cells. Also contemplated as an element of the expression cassette is a terminator. These elements can serve to enhance message levels and to minimize read through from the cassette into other sequences.

A specific initiation signal also may be required for efficient translation of coding sequences. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may need to be provided. One of ordinary skill in the art would readily be capable of determining this and providing the necessary signals. It is well known that the initiation codon must be "in-frame" with the reading frame of the desired coding sequence to ensure translation of the entire insert. The exogenous translational control signals and initiation codons can be either natural or synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements.

It is proposed that wild-type, polymorphic or mutant AH-genetically-associated genes may be co-expressed with other genes believed to be involved in bone loss or the hyperabsorption of calcium, such as 1-α hydroxylase, Vitamin D Receptor, calmodulin, IL-1α or β, IL-1α receptor or IL-1 β receptor, or parathyroid hormone-related proteins, wherein the proteins may be co-expressed in the same cell or wherein wild-type, polymorphic or mutant AH-genetically-associated genes may be provided to a cell that already has the other genes involved in bone loss or the hyperabsorption of calcium. Co-expression may be achieved by co-transfecting the cell with two distinct recombinant

vectors, each bearing a copy of either the respective DNA. Alternatively, a single recombinant vector may be constructed to include the coding regions for both of the proteins, which could then be expressed in cells transfected with the single vector. In either event, the term "co-expression" herein refers to the expression of both the wild-type, polymorphic or mutant AH-genetically-associated genes and the genes encoding proteins involved in bone loss or in hyperabsorption of calcium in the same recombinant cell.

As used herein, the terms "engineered" and "recombinant" cells are intended to refer to a cell into which an exogenous DNA segment or gene, such as a cDNA or gene encoding an AH-genetically-associated protein has been introduced. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced exogenous DNA segment or gene. Engineered cells are thus cells having a gene or genes introduced through the hand of man. Recombinant cells include those having an introduced cDNA or genomic gene, and also include genes positioned adjacent to a promoter not naturally associated with the particular introduced gene.

To express a recombinant AH-genetically-associated protein, whether mutant or wild-type, in accordance with the present invention one would prepare an expression vector that comprises a wild-type, polymorphic or mutant AH-genetically-associated nucleic acid under the control of one or more promoters. To bring a coding sequence "under the control of" a promoter, one positions the 5' end of the transcription initiation site of the transcriptional reading frame generally between about 1 and about 50 nucleotides "downstream" of (*i.e.*, 3' of) the chosen promoter. The "upstream" promoter stimulates transcription of the DNA and promotes expression of the encoded recombinant protein. This is the meaning of "recombinant expression" in this context.

Many standard techniques are available to construct expression vectors containing the appropriate nucleic acids and transcriptional/translational control sequences in order

to achieve protein or peptide expression in a variety of host-expression systems. Cell types available for expression include, but are not limited to, bacteria, such as *E. coli* and *B. subtilis* transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors.

Certain examples of prokaryotic hosts are *E. coli* strain RR1, *E. coli* LE392, *E. coli* B, *E. coli* X 1776 (ATCC No. 31537) as well as *E. coli* W3110 (F-, lambda-, prototrophic, ATCC No. 273325); bacilli such as *Bacillus subtilis*; and other enterobacteriaceae such as *Salmonella typhimurium*, *Serratia marcescens*, and various *Pseudomonas* species.

In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is often transformed using derivatives of pBR322, a plasmid derived from an *E. coli* species. pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism can be used as transforming vectors in connection with these hosts. For example, the phage lambda GEM™-11 may be utilized in making a recombinant phage vector which can be used to transform host cells, such as *E. coli* LE392.

Further useful vectors include pIN vectors (Inouye *et al.*, 1985); and pGEX vectors, for use in generating glutathione S-transferase (GST) soluble fusion proteins for

later purification and separation or cleavage. Other suitable fusion proteins are those with β-galactosidase, ubiquitin, and the like.

Promoters that are most commonly used in recombinant DNA construction include the β-lactamase (penicillinase), lactose and tryptophan (trp) promoter systems. While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling those of skill in the art to ligate them functionally with plasmid vectors.

The following details concerning recombinant protein production in bacterial cells, such as *E. coli*, are provided by way of exemplary information on recombinant protein production in general, the adaptation of which to a particular recombinant expression system will be known to those of skill in the art.

Bacterial cells, for example, *E. coli*, containing the expression vector are grown in any of a number of suitable media, for example, Luria Broth (LB). The expression of the recombinant protein may be induced, *e.g.*, by adding isopropyl-β-D-thiogalactoside (IPTG) to the media or by switching incubation to a higher temperature. After culturing the bacteria for a further period, generally of between 2 and 24 hours, the cells are collected by centrifugation and washed to remove residual media.

The bacterial cells are then lysed, for example, by disruption in a cell homogenizer and centrifuged to separate the dense inclusion bodies and cell membranes from the soluble cell components. This centrifugation can be performed under conditions whereby the dense inclusion bodies are selectively enriched by incorporation of sugars, such as sucrose, into the buffer and centrifugation at a selective speed.

If the recombinant protein is expressed in the inclusion bodies, as is the case in many instances, these can be washed in any of several solutions to remove some of the

contaminating host proteins, then solubilized in solutions containing high concentrations of urea (*e.g.* 8M) or chaotropic agents such as guanidine hydrochloride in the presence of reducing agents, such as β-mercaptoethanol or DTT (dithiothreitol).

Under some circumstances, it may be advantageous to incubate the protein for several hours under conditions suitable for the protein to undergo a refolding process into a conformation which more closely resembles that of the native protein. Such conditions generally include low protein concentrations, less than 500 mg/ml, low levels of reducing agent, concentrations of urea less than 2 M and often the presence of reagents such as a mixture of reduced and oxidized glutathione which facilitate the interchange of disulfide bonds within the protein molecule.

The refolding process can be monitored, for example, by SDS-PAGE, or with antibodies specific for the native molecule (which can be obtained from animals vaccinated with the native molecule or smaller quantities of recombinant protein). Following refolding, the protein can then be purified further and separated from the refolding mixture by chromatography on any of several supports including ion exchange resins, gel permeation resins or on a variety of affinity columns.

For expression in *Saccharomyces*, the plasmid YRp7, for example, is commonly used. This plasmid already contains the *trp*l gene which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1. The presence of the *trp*l lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase or other glycolytic enzymes, such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase,

triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination.

Other suitable promoters, which have the additional advantage of transcription controlled by growth conditions, include the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization.

In addition to micro-organisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture. In addition to mammalian cells, these include insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus); and plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing one or more wild-type, polymorphic or mutant AH-genetically-associated coding sequences.

In a useful insect system, *Autograph californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The wild-type, polymorphic or mutant AH-genetically-associated coding sequences are cloned into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the coding sequences results in the inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (*e.g.*, U.S. Patent No. 4,215,051, Smith, incorporated herein by reference).

Examples of useful mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, W138, BHK, COS-7, 293, HepG2, 3T3, RIN and MDCK cell lines. In addition, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the protein.

Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cells lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed.

Expression vectors for use in mammalian such cells ordinarily include an origin of replication (as necessary), a promoter located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences. The origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (*e.g.*, Polyoma, Adenovirus, VSV, BPV) source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

The promoters may be derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter). Further, it is also possible, and may be desirable, to utilize promoter or control sequences normally associated with the desired wild-type, polymorphic or mutant AH-genetically-associated gene sequence, provided such control sequences are compatible with the host cell systems.

A number of viral based expression systems may be utilized, for example, commonly used promoters are derived from polyoma, Adenovirus 2, and most frequently Simian Virus 40 (SV40). The early and late promoters of SV40 virus are particularly useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250 bp sequence extending from the *Hin*dIII site toward the *Bgl*I site located in the viral origin of replication.

In cases where an adenovirus is used as an expression vector, the coding sequences may be ligated to an adenovirus transcription/ translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing wild-type, polymorphic or mutant AH-genetically-associated proteins in infected hosts.

Specific initiation signals may also be required for efficient translation of wild-type, polymorphic or mutant AH-genetically-associated coding sequences. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may additionally need to be provided. One of ordinary skill in the art would readily be capable of determining this and providing the necessary signals. It is well known that the initiation codon must be in-frame (or in-phase) with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators.

In eukaryotic expression, one will also typically desire to incorporate into the transcriptional unit an appropriate polyadenylation site (*e.g.*, 5'-AATAAA-3') if one was

not contained within the original cloned segment. Typically, the poly A addition site is placed about 30 to 2000 nucleotides "downstream" of the termination site of the protein at a position prior to transcription termination.

For long-term, high-yield production of recombinant wild-type, polymorphic or mutant AH-genetically-associated proteins, stable expression is preferred. For example, cell lines that stably express constructs encoding wild-type, polymorphic or mutant AH-genetically-associated proteins may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with vectors controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines.

A number of selection systems may be used, including, but not limited, to the herpes simplex virus thymidine kinase, hypoxanthine-guanine phosphoribosyltransferase and adenine phosphoribosyltransferase genes, in tk-, hgprt- or aprt- cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, that confers resistance to methotrexate; gpt, that confers resistance to mycophenolic acid; neo, that confers resistance to the aminoglycoside G-418; and hygro, that confers resistance to hygromycin.

Animal cells can be propagated *in vitro* in two modes: as non-anchorage dependent cells growing in suspension throughout the bulk of the culture or as anchorage-dependent cells requiring attachment to a solid substrate for their propagation (*i.e.*, a monolayer type of cell growth).

Non-anchorage dependent or suspension cultures from continuous established cell lines are the most widely used means of large scale production of cells and cell products. However, suspension cultured cells have limitations, such as tumorigenic potential and lower protein production than adherent cells.

Large scale suspension culture of mammalian cells in stirred tanks is a common method for production of recombinant proteins. Two suspension culture reactor designs are in wide use - the stirred reactor and the airlift reactor. The stirred design has successfully been used on an 8000 liter capacity for the production of interferon. Cells are grown in a stainless steel tank with a height-to-diameter ratio of 1:1 to 3:1. The culture is usually mixed with one or more agitators, based on bladed disks or marine propeller patterns. Agitator systems offering less shear forces than blades have been described. Agitation may be driven either directly or indirectly by magnetically coupled drives. Indirect drives reduce the risk of microbial contamination through seals on stirrer shafts.

The airlift reactor, also initially described for microbial fermentation and later adapted for mammalian culture, relies on a gas stream to both mix and oxygenate the culture. The gas stream enters a riser section of the reactor and drives circulation. Gas disengages at the culture surface, causing denser liquid free of gas bubbles to travel downward in the downcomer section of the reactor. The main advantage of this design is the simplicity and lack of need for mechanical mixing. Typically, the height-to-diameter ratio is 10:1. The airlift reactor scales up relatively easily, has good mass transfer of gases and generates relatively low shear forces.

It is contemplated that the wild-type, polymorphic or mutant AH-genetically-associated proteins of the invention may be "overexpressed", *i.e.*, expressed in increased levels relative to its natural expression in cells. Such overexpression may be assessed by a variety of methods, including radio-labelling and/or protein purification. However, simple and direct methods are preferred, for example, those involving SDS/PAGE and

protein staining or western blotting, followed by quantitative analyses, such as densitometric scanning of the resultant gel or blot. A specific increase in the level of the recombinant protein or peptide in comparison to the level in natural cells is indicative of overexpression, as is a relative abundance of the specific protein in relation to the other proteins produced by the host cell and, *e.g.*, visible on a gel.

C. Nucleic Acid Detection

In addition to their use in directing the expression of the wild-type, polymorphic or mutant AH-genetically-associated proteins, the nucleic acid sequences disclosed herein also have a variety of other uses. For example, they also have utility as probes or primers in nucleic acid hybridization embodiments.

1. Hybridization

The use of a hybridization probe of between 17 and 100 nucleotides or longer in length allows the formation of a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 20 bases in length are generally preferred, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of particular hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having stretches of 20 to 30 nucleotides, or even longer where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means or by introducing selected sequences into recombinant vectors for recombinant production.

Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of genes or RNAs or to provide primers for amplification of DNA or RNA from tissues. Depending on the application envisioned, one will desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence.

For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.10 M NaCl at temperatures of about 50°C to about 70°C. Such high stringency conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating specific genes or detecting specific mRNA transcripts. It is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide.

For certain applications, for example, substitution of nucleotides by site-directed mutagenesis, it is appreciated that lower stringency conditions are required. Under these conditions, hybridization may occur even though the sequences of probe and target strand are not perfectly complementary, but are mismatched at one or more positions. Conditions may be rendered less stringent by increasing salt concentration and decreasing temperature. For example, a medium stringency condition could be provided by about 0.1 to 0.25 M NaCl at temperatures of about 37°C to about 55°C, while a low stringency condition could be provided by about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Thus, hybridization conditions can be readily manipulated depending on the desired results.

In other embodiments, hybridization may be achieved under conditions of, for example, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 1.0 mM dithiothreitol, at temperatures between approximately 20°C to about 37°C. Other hybridization conditions utilized could include approximately 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, at temperatures ranging from approximately 40°C to about 72°C.

In certain embodiments, it will be advantageous to employ nucleic acid sequences of the present invention in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin,

which are capable of being detected. In preferred embodiments, one may desire to employ a fluorescent label or an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmentally undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known that can be employed to provide a detection means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

In general, it is envisioned that the hybridization probes described herein will be useful both as reagents in solution hybridization, as in PCR, for detection of expression of corresponding genes, as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to hybridization with selected probes under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required (depending, for example, on the G+C content, type of target nucleic acid, source of nucleic acid, size of hybridization probe, *etc.*). Following washing of the hybridized surface to remove non-specifically bound probe molecules, hybridization is detected, or even quantified, by means of the label.

2. Amplification and PCR

Nucleic acid used as a template for amplification is isolated from cells contained in the biological sample, according to standard methodologies (Sambrook *et al.*, 1989). The nucleic acid may be genomic DNA or fractionated or whole cell RNA. Where RNA is used, it may be desired to convert the RNA to a complementary DNA. In one embodiment, the RNA is whole cell RNA and is used directly as the template for amplification.

Pairs of primers that selectively hybridize to nucleic acids corresponding to wild-type, polymorphic or mutant AH-genetically-associated protein are contacted with the isolated nucleic acid under conditions that permit selective hybridization. The term

"primer", as defined herein, is meant to encompass any nucleic acid that is capable of priming the synthesis of a nascent nucleic acid in a template-dependent process. Typically, primers are oligonucleotides from ten to twenty base pairs in length, but longer sequences can be employed. Primers may be provided in double-stranded or single-stranded form, although the single-stranded form is preferred.

Once hybridized, the nucleic acid:primer complex is contacted with one or more enzymes that facilitate template-dependent nucleic acid synthesis. Multiple rounds of amplification, also referred to as "cycles," are conducted until a sufficient amount of amplification product is produced.

Next, the amplification product is detected. In certain applications, the detection may be performed by visual means. Alternatively, the detection may involve indirect identification of the product via chemiluminescence, radioactive scintigraphy of incorporated radiolabel or fluorescent label or even via a system using electrical or thermal impulse signals (Affymax technology).

A number of template dependent processes are available to amplify the marker sequences present in a given template sample. One of the best known amplification methods is the polymerase chain reaction (referred to as PCR) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, and each incorporated herein by reference in entirety.

Briefly, to perform PCR, two primer sequences are prepared that are complementary to regions on opposite complementary strands of the marker sequence. An excess of deoxynucleoside triphosphates are added to a reaction mixture along with a DNA polymerase, *e.g.*, *Taq* polymerase. If the marker sequence is present in a sample, the primers will bind to the marker and the polymerase will cause the primers to be extended along the marker sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the

marker to form reaction products, excess primers will bind to the marker and to the reaction products and the process is repeated.

A reverse transcriptase PCR amplification procedure may be performed in order to quantify the amount of mRNA amplified. Methods of reverse transcribing RNA into cDNA are well known and described in Sambrook *et al.*, 1989. Alternative methods for reverse transcription utilize thermostable, RNA-dependent DNA polymerases. These methods are described in WO 90/07641, filed December 21, 1990, incorporated herein by reference. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction ("LCR"), disclosed in EPA No. 320 308, incorporated herein by reference in its entirety. In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent 4,883,750 describes a method similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Application No. PCT/US87/00880, incorporated herein by reference, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[alpha-thio]-triphosphates in one strand of a restriction site may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.*, nick translation. A similar method, called Repair Chain Reaction (RCR), involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA. Target specific sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having 3' and 5' sequences of non-specific DNA and a middle sequence of specific RNA is hybridized to DNA that is present in a sample. Upon hybridization, the reaction is treated with RNase H, and the products of the probe identified as distinctive products that are released after digestion. The original template is annealed to another cycling probe and the reaction is repeated.

Still another amplification methods described in GB Application No. 2 202 328, and in PCT Application No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template- and enzyme-dependent synthesis. The primers may be modified by labelling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes are added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS), including nucleic acid sequence based amplification (NASBA) and 3SR Gingeras *et al.*, PCT Application WO 88/10315, incorporated herein by reference. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a clinical sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride

extraction of RNA. These amplification techniques involve annealing a primer which has target specific sequences. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target specific primer, followed by polymerization. The double-stranded DNA molecules are then multiply transcribed by an RNA polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNA's are reverse transcribed into single stranded DNA, which is then converted to double stranded DNA, and then transcribed once again with an RNA polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target specific sequences.

Davey *et al.*, EPA No. 329 822 (incorporated herein by reference in its entirety) disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from the resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in duplex with either DNA or RNA). The resultant ssDNA is a template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to the template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting in a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

Miller *et al.*, PCT Application WO 89/06700 (incorporated herein by reference in its entirety) disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic, *i.e.*, new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" and "one-sided PCR" (Frohman, 1990 incorporated by reference).

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide, may also be used in the amplification step of the present invention.

Following any amplification, it may be desirable to separate the amplification product from the template and the excess primer for the purpose of determining whether specific amplification has occurred. In one embodiment, amplification products are separated by agarose, agarose-acrylamide or polyacrylamide gel electrophoresis using standard methods. See Sambrook *et al.*, 1989.

Alternatively, chromatographic techniques may be employed to effect separation. There are many kinds of chromatography which may be used in the present invention: adsorption, partition, ion-exchange and molecular sieve, and many specialized techniques for using them including column, paper, thin-layer and gas chromatography.

Amplification products must be visualized in order to confirm amplification of the marker sequences. One typical visualization method involves staining of a gel with ethidium bromide and visualization under UV light. Alternatively, if the amplification products are integrally labeled with radio- or fluorometrically-labeled nucleotides, the

amplification products can then be exposed to x-ray film or visualized under the appropriate stimulating spectra, following separation.

In one embodiment, visualization is achieved indirectly. Following separation of amplification products, a labeled, nucleic acid probe is brought into contact with the amplified marker sequence. The probe preferably is conjugated to a chromophore but may be radiolabeled. In another embodiment, the probe is conjugated to a binding partner, such as an antibody or biotin, and the other member of the binding pair carries a detectable moiety.

In one embodiment, detection is by Southern blotting and hybridization with a labeled probe. The techniques involved in Southern blotting are well known to those of skill in the art and can be found in many standard books on molecular protocols. See Sambrook *et al.*, 1989. Briefly, amplification products are separated by gel electrophoresis. The gel is then contacted with a membrane, such as nitrocellulose, permitting transfer of the nucleic acid and non-covalent binding. Subsequently, the membrane is incubated with a chromophore-conjugated probe that is capable of hybridizing with a target amplification product. Detection is by exposure of the membrane to x-ray film or ion-emitting detection devices.

One example of the foregoing is described in U.S. Patent No. 5,279,721, incorporated by reference herein, which discloses an apparatus and method for the automated electrophoresis and transfer of nucleic acids. The apparatus permits electrophoresis and blotting without external manipulation of the gel and is ideally suited to carrying out methods according to the present invention.

All the essential materials and reagents required for detecting wild-type, polymorphic or mutant AH-genetically-associated protein markers in a biological sample may be assembled together in a kit. This generally will comprise preselected primers for specific markers. Also included may be enzymes suitable for amplifying nucleic acids

including various polymerases (RT, Taq, *etc.*), deoxynucleotides and buffers to provide the necessary reaction mixture for amplification.

Such kits generally will comprise, in suitable means, distinct containers for each individual reagent and enzyme as well as for each marker primer pair. Preferred pairs of primers for amplifying nucleic acids are selected to amplify the sequences on human chromosome 1 that are genetically associated with AH such as the AH-genetically associated gene sequences specified in SEQ ID NO:1.

In another embodiment, such kits will comprise hybridization probes specific for wild-type, polymorphic or mutant AH-genetically-associated protein chosen from a group including nucleic acids corresponding to the sequences specified in SEQ ID NO:1. Such kits generally will comprise, in suitable means, distinct containers for each individual reagent and enzyme as well as for each marker hybridization probe.

3. Other Assays

Other methods for genetic screening to accurately detect mutations in genomic DNA, cDNA or RNA samples may be employed, depending on the specific situation.

Historically, a number of different methods have been used to detect point mutations, including denaturing gradient gel electrophoresis ("DGGE"), restriction enzyme polymorphism analysis, chemical and enzymatic cleavage methods, and others. The more common procedures currently in use include direct sequencing of target regions amplified by PCR™ (see above) and single-strand conformation polymorphism analysis ("SSCP").

Another method of screening for point mutations is based on RNase cleavage of base pair mismatches in RNA/DNA and RNA/RNA heteroduplexes. As used herein, the term "mismatch" is defined as a region of one or more unpaired or mispaired nucleotides in a double-stranded RNA/RNA, RNA/DNA or DNA/DNA molecule. This definition

thus includes mismatches due to insertion/deletion mutations, as well as single and multiple base point mutations.

U.S. Patent No. 4,946,773 describes an RNase A mismatch cleavage assay that involves annealing single-stranded DNA or RNA test samples to an RNA probe, and subsequent treatment of the nucleic acid duplexes with RNase A. After the RNase cleavage reaction, the RNase is inactivated by proteolytic digestion and organic extraction, and the cleavage products are denatured by heating and analyzed by electrophoresis on denaturing polyacrylamide gels. For the detection of mismatches, the single-stranded products of the RNase A treatment, electrophoretically separated according to size, are compared to similarly treated control duplexes. Samples containing smaller fragments (cleavage products) not seen in the control duplex are scored as +.

Currently available RNase mismatch cleavage assays, including those performed according to U.S. Patent No. 4,946,773, require the use of radiolabeled RNA probes. Myers and Maniatis in U.S. Patent No. 4,946,773 describe the detection of base pair mismatches using RNase A. Other investigators have described the use of *E. coli* enzyme, RNase I, in mismatch assays. Because it has broader cleavage specificity than RNase A, RNase I would be a desirable enzyme to employ in the detection of base pair mismatches if components can be found to decrease the extent of non-specific cleavage and increase the frequency of cleavage of mismatches. The use of RNase I for mismatch detection is described in literature from Promega Biotech. Promega markets a kit containing RNase I that is shown in their literature to cleave three out of four known mismatches, provided the enzyme level is sufficiently high.

The RNase protection assay was first used to detect and map the ends of specific mRNA targets in solution. The assay relies on being able to easily generate high specific activity radiolabeled RNA probes complementary to the mRNA of interest by *in vitro* transcription. Originally, the templates for *in vitro* transcription were recombinant plasmids containing bacteriophage promoters. The probes are mixed with total cellular

RNA samples to permit hybridization to their complementary targets, then the mixture is treated with RNase to degrade excess unhybridized probe. Also, as originally intended, the RNase used is specific for single-stranded RNA, so that hybridized double-stranded probe is protected from degradation. After inactivation and removal of the RNase, the protected probe (which is proportional in amount to the amount of target mRNA that was present) is recovered and analyzed on a polyacrylamide gel.

The RNase Protection assay was adapted for detection of single base mutations. In this type of RNase A mismatch cleavage assay, radiolabeled RNA probes transcribed *in vitro* from wild-type sequences, are hybridized to complementary target regions derived from test samples. The test target generally comprises DNA (either genomic DNA or DNA amplified by cloning in plasmids or by PCR™), although RNA targets (endogenous mRNA) have occasionally been used. If single nucleotide (or greater) sequence differences occur between the hybridized probe and target, the resulting disruption in Watson-Crick hydrogen bonding at that position ("mismatch") can be recognized and cleaved in some cases by single-strand specific ribonuclease. To date, RNase A has been used almost exclusively for cleavage of single-base mismatches, although RNase I has recently been shown as useful also for mismatch cleavage. There are recent descriptions of using the MutS protein and other DNA-repair enzymes for detection of single-base mismatches.

D. Mutagenesis

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent proteins or peptides, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of

sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is preferred, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art. As will be appreciated, the technique typically employs a bacteriophage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage vectors are commercially available and their use is generally well known to those skilled in the art. Double stranded plasmids are also routinely employed in site directed mutagenesis, which eliminates the step of transferring the gene of interest from a phage to a plasmid.

In general, site-directed mutagenesis is performed by first obtaining a single-stranded vector, or melting of two strands of a double stranded vector which includes within its sequence a DNA sequence encoding the desired protein. An oligonucleotide primer bearing the desired mutated sequence is synthetically prepared. This primer is then annealed with the single-stranded DNA preparation, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected that include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected gene using site-directed mutagenesis is provided as a means of producing potentially useful species and is not meant to be limiting, as there are other ways in which sequence variants of genes may be obtained. For example, recombinant vectors encoding the desired gene may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants.

II. AH-Genetically-Associated Proteins and Peptides

The present invention therefore provides purified, and in preferred embodiments, substantially purified, AH-genetically-associated proteins and peptides. The term "purified AH-genetically-associated protein or peptide" as used herein, is intended to refer to a wild-type, polymorphic or mutant AH-genetically-associated proteinaceous composition, isolatable from mammalian cells or recombinant host cells, wherein the wild-type, polymorphic or mutant AH-genetically-associated protein or peptide is purified to any degree relative to its naturally-obtainable state, *i.e.*, relative to its purity within a cellular extract. A purified wild-type, polymorphic or AH-genetically-associated protein or peptide therefore also refers to a wild-type, polymorphic or AH-genetically-associated protein or peptide free from the environment in which it naturally occurs. An AH-genetically associated protein or peptide is given as SEQ ID NO:2.

Wild-type, polymorphic or mutant AH-genetically-associated proteins may be full length proteins. Wild-type, polymorphic or mutant AH-genetically-associated proteins, polypeptides and peptides may also be less then full length proteins, such as individual domains, regions or even epitopic peptides. Where less than full length wild-type, polymorphic or mutant AH-genetically-associated proteins are concerned the most preferred will be those containing predicted immunogenic sites and those containing the functional domains identified herein.

Generally, "purified" will refer to a wild-type, polymorphic or AH-genetically-associated protein or peptide composition that has been subjected to fractionation to remove various non-wild-type, polymorphic or mutant AH-genetically-associated protein or peptide components, and which composition substantially retains its wild-type, polymorphic or mutant activity.

Where the term "substantially purified" is used, this will refer to a composition in which the wild-type, polymorphic or mutant AH-genetically-associated protein or peptide

forms the major component of the composition, such as constituting about 50% of the proteins in the composition or more. In preferred embodiments, a substantially purified protein will constitute more than 60%, 70%, 80%, 90%, 95%, 99% or even more of the proteins in the composition.

A polypeptide or protein that is "purified to homogeneity," as applied to the present invention, means that the polypeptide or protein has a level of purity where the polypeptide or protein is substantially free from other proteins and biological components. For example, a purified polypeptide or protein will often be sufficiently free of other protein components so that degradative sequencing may be performed successfully.

Various methods for quantifying the degree of purification of wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides will be known to those of skill in the art in light of the present disclosure. These include, for example, determining the specific activity of a fraction, or assessing the number of polypeptides within a fraction by gel electrophoresis. Assessing the number of polypeptides within a fraction by SDS/PAGE analysis will often be preferred in the context of the present invention as this is straightforward.

To purify a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide a natural or recombinant composition comprising at least some wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides will be subjected to fractionation to remove various non-wild-type, polymorphic or AH-genetically-associated components from the composition. Various techniques suitable for use in protein purification will be well known to those of skill in the art. These include, for example, precipitation with ammonium sulfate, PEG, antibodies and the like or by heat denaturation, followed by centrifugation; chromatography steps such as ion exchange, gel filtration, reverse phase, hydroxylapatite, lectin affinity and other affinity chromatography

steps; isoelectric focusing; gel electrophoresis; and combinations of such and other techniques.

Another purification method involves a fusion protein using a specific binding partner. Such purification methods are routine in the art. As the present invention provides DNA sequences for AH-genetically-associated proteins, any fusion protein purification method can now be practiced. This is currently exemplified by the generation of a AH-genetically-associated-glutathione S-transferase fusion protein, expression in *E. coli*, and isolation to homogeneity using affinity chromatography on glutathione-agarose.

The exemplary purification method disclosed herein represents one method to prepare a substantially purified wild-type, polymorphic or mutant AH-genetically-associated protein or peptide. This method is preferred as it results in the substantial purification of the wild-type, polymorphic or mutant AH-genetically-associated protein or peptide in yields sufficient for further characterization and use. However, given the DNA and proteins provided by the present invention, any purification method can now be employed.

Although preferred for use in certain embodiments, there is no general requirement that the wild-type, polymorphic or mutant AH-genetically-associated protein or peptide always be provided in their most purified state. Indeed, it is contemplated that less substantially purified wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides, which are nonetheless enriched in wild-type, polymorphic or mutant AH-genetically-associated protein compositions, relative to the natural state, will have utility in certain embodiments. These include, for example antibody generation where subsequent screening assays using purified wild-type, polymorphic or AH-genetically-associated proteins are conducted.

Methods exhibiting a lower degree of relative purification may have advantages in total recovery of protein product, or in maintaining the activity of an expressed protein. Inactive products also have utility in certain embodiments, such as, *e.g.*, in antibody generation.

III. Antibodies to AH-genetically-associated Proteins

A. Epitopic Core Sequences

Peptides corresponding to one or more antigenic determinants, or "epitopic core regions", of wild-type, polymorphic or mutant AH-genetically-associated proteins of the present invention can also be prepared. Such peptides should generally be at least five or six amino acid residues in length, will preferably be about 10, 15, 20, 25 or about 30 amino acid residues in length, and may contain up to about 35-50 residues or so.

Synthetic peptides will generally be about 35 residues long, which is the approximate upper length limit of automated peptide synthesis machines, such as those available from Applied Biosystems (Foster City, CA). Longer peptides may also be prepared, *e.g.*, by recombinant means.

U.S. Patent 4,554,101, (Hopp) incorporated herein by reference, teaches the identification and preparation of epitopes from primary amino acid sequences on the basis of hydrophilicity. Through the methods disclosed in Hopp, one of skill in the art would be able to identify epitopes from within an amino acid sequence such as the wild-type, polymorphic or mutant AH-genetically-associated sequences disclosed herein.

Numerous scientific publications have also been devoted to the prediction of secondary structure, and to the identification of epitopes, from analyses of amino acid sequences (Chou & Fasman, 1974a,b; 1978a,b, 1979). Any of these may be used, if desired, to supplement the teachings of Hopp in U.S. Patent 4,554,101.

Moreover, computer programs are currently available to assist with predicting antigenic portions and epitopic core regions of proteins. Examples include those programs based upon the Jameson-Wolf analysis (Jameson & Wolf, 1998; Wolf *et al.*, 1988), the program PepPlot® (Brutlag *et al.*, 1990; Weinberger *et al.*, 1985), and other new programs for protein tertiary structure prediction (Fetrow & Bryant, 1993). Further commercially available software capable of carrying out such analyses is termed MacVector (IBI, New Haven, CT).

In further embodiments, major antigenic determinants of a polypeptide may be identified by an empirical approach in which portions of the gene encoding the polypeptide are expressed in a recombinant host, and the resulting proteins tested for their ability to elicit an immune response. For example, PCR can be used to prepare a range of peptides lacking successively longer fragments of the C-terminus of the protein. The immunoactivity of each of these peptides is determined to identify those fragments or domains of the polypeptide that are immunodominant. Further studies in which only a small number of amino acids are removed at each iteration then allows the location of the antigenic determinants of the polypeptide to be more precisely determined.

Another method for determining the major antigenic determinants of a polypeptide is the SPOTs™ system (Genosys Biotechnologies, Inc., The Woodlands, TX). In this method, overlapping peptides are synthesized on a cellulose membrane, which following synthesis and deprotection, is screened using a polyclonal or monoclonal antibody. The antigenic determinants of the peptides which are initially identified can be further localized by performing subsequent syntheses of smaller peptides with larger overlaps, and by eventually replacing individual amino acids at each position along the immunoreactive peptide.

Once one or more such analyses are completed, polypeptides are prepared that contain at least the essential features of one or more antigenic determinants. The peptides are then employed in the generation of antisera against the polypeptide. Minigenes or

gene fusions encoding these determinants can also be constructed and inserted into expression vectors by standard methods, for example, using PCR cloning methodology.

The use of such small peptides for vaccination typically requires conjugation of the peptide to an immunogenic carrier protein, such as hepatitis B surface antigen, keyhole limpet hemocyanin or bovine serum albumin. Methods for performing this conjugation are well known in the art.

B. Antibody Generation

In certain embodiments, the present invention provides antibodies that bind with high specificity to wild-type, polymorphic or mutant AH-genetically-associated proteins provided herein. Thus, antibodies that bind to the protein products of the isolated nucleic acid sequences of SEQ ID NO:1 are provided. Antibodies specific for the wild-type and polymorphic proteins and peptides and those specific for any one of a number of mutants are provided. As detailed above, in addition to antibodies generated against the full length proteins, antibodies may also be generated in response to smaller constructs comprising epitopic core regions, including wild-type, polymorphic and mutant epitopes.

As used herein, the term "antibody" is intended to refer broadly to any immunologic binding agent such as IgG, IgM, IgA, IgD and IgE. Generally, IgG and/or IgM are preferred because they are the most common antibodies in the physiological situation and because they are most easily made in a laboratory setting.

Monoclonal antibodies (MAbs) are recognized to have certain advantages, *e.g.*, reproducibility and large-scale production, and their use is generally preferred. The invention thus provides monoclonal antibodies of the human, murine, monkey, rat, hamster, rabbit and even chicken origin. Due to the ease of preparation and ready availability of reagents, murine monoclonal antibodies will often be preferred.

However, "humanized" antibodies are also contemplated, as are chimeric antibodies from mouse, rat, or other species, bearing human constant and/or variable region domains, bispecific antibodies, recombinant and engineered antibodies and fragments thereof. Methods for the development of antibodies that are "custom-tailored" to the patient's tumor are likewise known and such custom-tailored antibodies are also contemplated.

The term "antibody" is used to refer to any antibody-like molecule that has an antigen binding region, and includes antibody fragments such as Fab', Fab, $F(ab')_2$, single domain antibodies (DABs), Fv, scFv (single chain Fv), and the like. The techniques for preparing and using various antibody-based constructs and fragments are well known in the art.

Means for preparing and characterizing antibodies are well known in the art (See, *e.g.*, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988; incorporated herein by reference).

The methods for generating monoclonal antibodies (MAbs) generally begin along the same lines as those for preparing polyclonal antibodies. Briefly, a polyclonal antibody is prepared by immunizing an animal with an immunogenic wild-type, polymorphic or mutant AH-genetically-associated protein composition in accordance with the present invention and collecting antisera from that immunized animal.

A wide range of animal species can be used for the production of antisera. Typically the animal used for production of anti-antisera is a rabbit, a mouse, a rat, a hamster, a guinea pig or a goat. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies.

As is well known in the art, a given composition may vary in its immunogenicity. It is often necessary therefore to boost the host immune system, as may be achieved by

coupling a peptide or polypeptide immunogen to a carrier. Exemplary and preferred carriers are keyhole limpet hemocyanin (KLH) and bovine serum albumin (BSA). Other albumins such as ovalbumin, mouse serum albumin or rabbit serum albumin can also be used as carriers. Means for conjugating a polypeptide to a carrier protein are well known in the art and include glutaraldehyde, m-maleimidobenzoyl-N-hydroxysuccinimide ester, carbodiimide and bis-biazotized benzidine.

As is also well known in the art, the immunogenicity of a particular immunogen composition can be enhanced by the use of non-specific stimulators of the immune response, known as adjuvants. Suitable adjuvants include all acceptable immunostimulatory compounds, such as cytokines, toxins or synthetic compositions.

Adjuvants that may be used include IL-1, IL-2, IL-4, IL-7, IL-12, g-interferon, GMCSP, BCG, aluminum hydroxide, MDP compounds, such as thur-MDP and nor-MDP, CGP (MTP-PE), lipid A, and monophosphoryl lipid A (MPL). RIBI, which contains three components extracted from bacteria, MPL, trehalose dimycolate (TDM) and cell wall skeleton (CWS) in a 2% squalene/Tween 80 emulsion. MHC antigens may even be used.

Exemplary, often preferred adjuvants include complete Freund's adjuvant (a non-specific stimulator of the immune response containing killed *Mycobacterium tuberculosis)*, incomplete Freund's adjuvants and aluminum hydroxide adjuvant.

In addition to adjuvants, it may be desirable to coadminister biologic response modifiers (BRM), which have been shown to upregulate T cell immunity or downregulate suppressor cell activity. Such BRMs include, but are not limited to, Cimetidine (CIM; 1200 mg/d) (Smith/Kline, PA); or low-dose Cyclophosphamide (CYP; 300 mg/m^2) (Johnson/ Mead, NJ) and Cytokines such as γ-interferon, IL-2, or IL-12 or genes encoding proteins involved in immune helper functions, such as B-7.

The amount of immunogen composition used in the production of polyclonal antibodies varies upon the nature of the immunogen as well as the animal used for immunization. A variety of routes can be used to administer the immunogen (subcutaneous, intramuscular, intradermal, intravenous and intraperitoneal). The production of polyclonal antibodies may be monitored by sampling blood of the immunized animal at various points following immunization.

A second, booster injection, may also be given. The process of boosting and titering is repeated until a suitable titer is achieved. When a desired level of immunogenicity is obtained, the immunized animal can be bled and the serum isolated and stored, and/or the animal can be used to generate MAbs.

For production of rabbit polyclonal antibodies, the animal can be bled through an ear vein or alternatively by cardiac puncture. The removed blood is allowed to coagulate and then centrifuged to separate serum components from whole cells and blood clots. The serum may be used as is for various applications or else the desired antibody fraction may be purified by well-known methods, such as affinity chromatography using another antibody, a peptide bound to a solid matrix, or by using, *e.g.*, protein A or protein G chromatography.

MAbs may be readily prepared through use of well-known techniques, such as those exemplified in U.S. Patent 4,196,265, incorporated herein by reference. Typically, this technique involves immunizing a suitable animal with a selected immunogen composition, *e.g.*, a purified or partially purified wild-type, polymorphic or mutant AH-genetically-associated protein, polypeptide, peptide or domain, be it a wild-type or mutant composition. The immunizing composition is administered in a manner effective to stimulate antibody producing cells.

The methods for generating monoclonal antibodies (MAbs) generally begin along the same lines as those for preparing polyclonal antibodies. Rodents such as mice and

rats are preferred animals, however, the use of rabbit, sheep frog cells is also possible. The use of rats may provide certain advantages (Goding, 1986, pp. 60-61), but mice are preferred, with the BALB/c mouse being most preferred as this is most routinely used and generally gives a higher percentage of stable fusions.

The animals are injected with antigen, generally as described above. The antigen may be coupled to carrier molecules such as keyhole limpet hemocyanin if necessary. The antigen would typically be mixed with adjuvant, such as Freund's complete or incomplete adjuvant. Booster injections with the same antigen would occur at approximately two-week intervals.

Following immunization, somatic cells with the potential for producing antibodies, specifically B lymphocytes (B cells), are selected for use in the MAb generating protocol. These cells may be obtained from biopsied spleens, tonsils or lymph nodes, or from a peripheral blood sample. Spleen cells and peripheral blood cells are preferred, the former because they are a rich source of antibody-producing cells that are in the dividing plasmablast stage, and the latter because peripheral blood is easily accessible.

Often, a panel of animals will have been immunized and the spleen of animal with the highest antibody titer will be removed and the spleen lymphocytes obtained by homogenizing the spleen with a syringe. Typically, a spleen from an immunized mouse contains approximately 5×10^7 to 2×10^8 lymphocytes.

The antibody-producing B lymphocytes from the immunized animal are then fused with cells of an immortal myeloma cell, generally one of the same species as the animal that was immunized. Myeloma cell lines suited for use in hybridoma-producing fusion procedures preferably are non-antibody-producing, have high fusion efficiency, and enzyme deficiencies that render then incapable of growing in certain selective media which support the growth of only the desired fused cells (hybridomas).

Any one of a number of myeloma cells may be used, as are known to those of skill in the art (Goding, pp. 65-66, 1986; Campbell, pp. 75-83, 1984). For example, where the immunized animal is a mouse, one may use P3-X63/Ag8, X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 and S194/5XX0 Bul; for rats, one may use R210.RCY3, Y3-Ag 1.2.3, IR983F and 4B210; and U-266, GM1500-GRG2, LICR-LON-HMy2 and UC729-6 are all useful in connection with human cell fusions.

One preferred murine myeloma cell is the NS-1 myeloma cell line (also termed P3-NS-1-Ag4-1), which is readily available from the NIGMS Human Genetic Mutant Cell Repository by requesting cell line repository number GM3573. Another mouse myeloma cell line that may be used is the 8-azaguanine-resistant mouse murine myeloma SP2/0 non-producer cell line.

Methods for generating hybrids of antibody-producing spleen or lymph node cells and myeloma cells usually comprise mixing somatic cells with myeloma cells in a 2:1 proportion, though the proportion may vary from about 20:1 to about 1:1, respectively, in the presence of an agent or agents (chemical or electrical) that promote the fusion of cell membranes. Fusion methods using Sendai virus have been described by Kohler and Milstein (1975; 1976), and those using polyethylene glycol (PEG), such as 37% (v/v) PEG, by Gefter *et al.* (1977). The use of electrically induced fusion methods is also appropriate (Goding pp. 71-74, 1986).

Fusion procedures usually produce viable hybrids at low frequencies, about 1×10^{-6} to 1×10^{-8}. However, this does not pose a problem, as the viable, fused hybrids are differentiated from the parental, unfused cells (particularly the unfused myeloma cells that would normally continue to divide indefinitely) by culturing in a selective medium. The selective medium is generally one that contains an agent that blocks the *de novo* synthesis of nucleotides in the tissue culture media. Exemplary and preferred agents are aminopterin, methotrexate, and azaserine. Aminopterin and methotrexate block *de novo* synthesis of both purines and pyrimidines, whereas azaserine blocks only purine

synthesis. Where aminopterin or methotrexate is used, the media is supplemented with hypoxanthine and thymidine as a source of nucleotides (HAT medium). Where azaserine is used, the media is supplemented with hypoxanthine.

The preferred selection medium is HAT. Only cells capable of operating nucleotide salvage pathways are able to survive in HAT medium. The myeloma cells are defective in key enzymes of the salvage pathway, *e.g.*, hypoxanthine phosphoribosyl transferase (HPRT), and they cannot survive. The B cells can operate this pathway, but they have a limited life span in culture and generally die within about two weeks. Therefore, the only cells that can survive in the selective media are those hybrids formed from myeloma and B cells.

This culturing provides a population of hybridomas from which specific hybridomas are selected. Typically, selection of hybridomas is performed by culturing the cells by single-clone dilution in microtiter plates, followed by testing the individual clonal supernatants (after about two to three weeks) for the desired reactivity. The assay should be sensitive, simple and rapid, such as radioimmunoassays, enzyme immunoassays, cytotoxicity assays, plaque assays, dot immunobinding assays, and the like.

The selected hybridomas would then be serially diluted and cloned into individual antibody-producing cell lines, which clones can then be propagated indefinitely to provide MAbs. The cell lines may be exploited for MAb production in two basic ways.

A sample of the hybridoma can be injected (often into the peritoneal cavity) into a histocompatible animal of the type that was used to provide the somatic and myeloma cells for the original fusion (*e.g.*, a syngeneic mouse). Optionally, the animals are primed with a hydrocarbon, especially oils such as pristane (tetramethylpentadecane) prior to injection. The injected animal develops tumors secreting the specific monoclonal

antibody produced by the fused cell hybrid. The body fluids of the animal, such as serum or ascites fluid, can then be tapped to provide MAbs in high concentration.

The individual cell lines could also be cultured *in vitro*, where the MAbs are naturally secreted into the culture medium from which they can be readily obtained in high concentrations.

MAbs produced by either means may be further purified, if desired, using filtration, centrifugation and various chromatographic methods such as HPLC or affinity chromatography. Fragments of the monoclonal antibodies of the invention can be obtained from the monoclonal antibodies so produced by methods which include digestion with enzymes, such as pepsin or papain, and/or by cleavage of disulfide bonds by chemical reduction. Alternatively, monoclonal antibody fragments encompassed by the present invention can be synthesized using an automated peptide synthesizer.

It is also contemplated that a molecular cloning approach may be used to generate monoclonals. For this, combinatorial immunoglobulin phagemid libraries are prepared from RNA isolated from the spleen of the immunized animal, and phagemids expressing appropriate antibodies are selected by panning using cells expressing the antigen and control cells. The advantages of this approach over conventional hybridoma techniques are that approximately 10^4 times as many antibodies can be produced and screened in a single round, and that new specificities are generated by H and L chain combination which further increases the chance of finding appropriate antibodies.

Alternatively, monoclonal antibody fragments encompassed by the present invention can be synthesized using an automated peptide synthesizer, or by expression of full-length gene or of gene fragments in *E. coli*.

C. Antibody Conjugates

The present invention further provides antibodies against wild-type, polymorphic or mutant AH-genetically-associated proteins, generally of the monoclonal type, that are linked to one or more other agents to form an antibody conjugate. Any antibody of sufficient selectivity, specificity and affinity may be employed as the basis for an antibody conjugate. Such properties may be evaluated using conventional immunological screening methodology known to those of skill in the art.

Certain examples of antibody conjugates are those conjugates in which the antibody is linked to a detectable label. "Detectable labels" are compounds or elements that can be detected due to their specific functional properties, or chemical characteristics, the use of which allows the antibody to which they are attached to be detected, and further quantified if desired. Another such example is the formation of a conjugate comprising an antibody linked to a cytotoxic or anti-cellular agent, as may be termed "immunotoxins". In the context of the present invention, immunotoxins are generally less preferred.

Antibody conjugates are thus preferred for use as diagnostic agents. Antibody diagnostics generally fall within two classes, those for use in *in vitro* diagnostics, such as in a variety of immunoassays, and those for use *in vivo* diagnostic protocols, generally known as "antibody-directed imaging". Again, antibody-directed imaging is less preferred for use with this invention.

Many appropriate imaging agents are known in the art, as are methods for their attachment to antibodies (see, *e.g.*, U.S. patents 5,021,236 and 4,472,509, both incorporated herein by reference). Certain attachment methods involve the use of a metal chelate complex employing, for example, an organic chelating agent such a DTPA attached to the antibody (U.S. Patent 4,472,509). Monoclonal antibodies may also be reacted with an enzyme in the presence of a coupling agent such as glutaraldehyde or

periodate. Conjugates with fluorescein markers are prepared in the presence of these coupling agents or by reaction with an isothiocyanate.

In the case of paramagnetic ions, one might mention by way of example ions such as chromium (III), manganese (II), iron (III), iron (II), cobalt (II), nickel (II), copper (II), neodymium (III), samarium (III), ytterbium (III), gadolinium (III), vanadium (II), terbium (III), dysprosium (III), holmium (III) and erbium (III), with gadolinium being particularly preferred.

Ions useful in other contexts, such as X-ray imaging, include but are not limited to lanthanum (III), gold (III), lead (II), and especially bismuth (III).

In the case of radioactive isotopes for therapeutic and/or diagnostic application, one might mention ^{211}astatine, ^{14}carbon, ^{51}chromium, ^{36}chlorine, ^{57}cobalt, ^{58}cobalt, ^{67}copper, ^{152}Eu, ^{67}gallium, ^{3}hydrogen, ^{123}iodine, ^{125}iodine, ^{131}iodine, ^{111}indium, ^{59}iron, ^{32}phosphorus, ^{186}rhenium, ^{188}rhenium, ^{75}selenium, ^{35}sulphur, ^{99m}technicium and ^{90}yttrium. ^{125}I is often being preferred for use in certain embodiments, and ^{99m}technicium and ^{111}indium are also often preferred due to their low energy and suitability for long range detection.

Radioactively labeled monoclonal antibodies of the present invention may be produced according to well-known methods in the art. For instance, monoclonal antibodies can be iodinated by contact with sodium or potassium iodide and a chemical oxidizing agent such as sodium hypochlorite, or an enzymatic oxidizing agent, such as lactoperoxidase. Monoclonal antibodies according to the invention may be labeled with ^{99m}technicium by ligand exchange process, for example, by reducing pertechnate with stannous solution, chelating the reduced technetium onto a Sephadex column and applying the antibody to this column or by direct labeling techniques, *e.g.*, by incubating pertechnate, a reducing agent such as $SNCl_2$, a buffer solution such as sodium-potassium phthalate solution, and the antibody.

Intermediary functional groups which are often used to bind radioisotopes which exist as metallic ions to antibody are diethylenetriaminepentaacetic acid (DTPA) and ethylene diaminetetracetic acid (EDTA).

Fluorescent labels include rhodamine, fluorescein isothiocyanate and renographin.

The much preferred antibody conjugates of the present invention are those intended primarily for use *in vitro*, where the antibody is linked to a secondary binding ligand or to an enzyme (an enzyme tag) that will generate a colored product upon contact with a chromogenic substrate. Examples of suitable enzymes include urease, alkaline phosphatase, (horseradish) hydrogen peroxidase and glucose oxidase. Preferred secondary binding ligands are biotin and avidin or streptavidin compounds. The use of such labels is well known to those of skill in the art in light and is described, for example, in U.S. Patents 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241; each incorporated herein by reference.

D. Immunodetection Methods

In still further embodiments, the present invention concerns immunodetection methods for binding, purifying, removing, quantifying or otherwise generally detecting biological components such as wild-type, polymorphic or mutant AH-genetically-associated protein components. The wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides of the present invention may be employed to detect and purify AH-genetically-associated polypeptides, and antibodies prepared in accordance with the present invention, may be employed to detect wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides. As described throughout the present application, the use of wild-type, polymorphic and mutant specific antibodies is contemplated. The steps of various useful immunodetection methods have been described in the scientific literature, such as, *e.g.*, Nakamura *et al.* (1987), incorporated herein by reference.

In general, the immunobinding methods include obtaining a sample suspected of containing a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide, and contacting the sample with a first anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibody in accordance with the present invention, as the case may be, under conditions effective to allow the formation of immunocomplexes.

These methods include methods for purifying wild-type, polymorphic or mutant AH-genetically-associated protein, as may be employed in purifiying wild-type, polymorphic or mutant AH-genetically-associated protein from patients' samples or for purifying recombinantly expressed wild-type, polymorphic or mutant AH-genetically-associated protein. In these instances, the antibody removes the antigenic wild-type, polymorphic or mutant AH-genetically-associated protein component from a sample. The antibody will preferably be linked to a solid support, such as in the form of a column matrix, and the sample suspected of containing the wild-type, polymorphic or mutant AH-genetically-associated protein antigenic component will be applied to the immobilized antibody. The unwanted components will be washed from the column, leaving the antigen immunocomplexed to the immobilized antibody, which wild-type, polymorphic or mutant AH-genetically-associated protein antigen is then collected by removing the wild-type, polymorphic or mutant AH-genetically-associated protein from the column.

The immunobinding methods also include methods for detecting or quantifying the amount of a wild-type, polymorphic or mutant AH-genetically-associated protein reactive component in a sample, which methods require the detection or quantification of any immune complexes formed during the binding process. Here, one would obtain a sample suspected of containing a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide, and contact the sample with an antibody against wild-type, polymorphic or mutant AH-genetically-associated protein, and then detect or quantify the amount of immune complexes formed under the specific conditions.

In terms of antigen detection, the biological sample analyzed may be any sample that is suspected of containing a wild-type, polymorphic or mutant AH-genetically-associated protein-specific antigen, such as a blood or urinary sample or specimen, a bone or kidney tissue section or specimen, a homogenized bone or kidney tissue extract, a bone or kidney cell, separated or purified forms of any of the above wild-type, polymorphic or mutant AH-genetically-associated protein-containing compositions, or even any biological fluid that comes into contact with bone or kidney tissue, including blood and serum.

Contacting the chosen biological sample with the antibody under conditions effective and for a period of time sufficient to allow the formation of immune complexes (primary immune complexes) is generally a matter of simply adding the antibody composition to the sample and incubating the mixture for a period of time lone enough for the antibodies to form immune complexes with, *i.e.*, to bind to, any wild-type, polymorphic or mutant AH-genetically-associated protein antigens present. After this time, the sample-antibody composition, such as a tissue section, ELISA plate, dot blot or western blot, will generally be washed to remove any non-specifically bound antibody species, allowing only those antibodies specifically bound within the primary immune complexes to be detected.

In general, the detection of immunocomplex formation is well known in the art and may be achieved through the application of numerous approaches. These methods are generally based upon the detection of a label or marker, such as any of those radioactive, fluorescent, biological or enzymatic tags. U.S. Patents concerning the use of such labels include 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241, each incorporated herein by reference. Of course, one may find additional advantages through the use of a secondary binding ligand such as a second antibody or a biotin/avidin ligand binding arrangement, as is known in the art.

The wild-type, polymorphic or mutant AH-genetically-associated protein antibody employed in the detection may itself be linked to a detectable label, wherein one would then simply detect this label, thereby allowing the amount of the primary immune complexes in the composition to be determined.

Alternatively, the first antibody that becomes bound within the primary immune complexes may be detected by means of a second binding ligand that has binding affinity for the antibody. In these cases, the second binding ligand may be linked to a detectable label. The second binding ligand is itself often an antibody, which may thus be termed a "secondary" antibody. The primary immune complexes are contacted with the labeled, secondary binding ligand, or antibody, under conditions effective and for a period of time sufficient to allow the formation of secondary immune complexes. The secondary immune complexes are then generally washed to remove any non-specifically bound labeled secondary antibodies or ligands, and the remaining label in the secondary immune complexes is then detected.

Further methods include the detection of primary immune complexes by a two step approach. A second binding ligand, such as an antibody, that has binding affinity for the antibody is used to form secondary immune complexes, as described above. After washing, the secondary immune complexes are contacted with a third binding ligand or antibody that has binding affinity for the second antibody, again under conditions effective and for a period of time sufficient to allow the formation of immune complexes (tertiary immune complexes). The third ligand or antibody is linked to a detectable label, allowing detection of the tertiary immune complexes thus formed. This system may provide for signal amplification if this is desired.

The immunodetection methods of the present invention have evident utility in the diagnosis or prognosis of conditions such stone formation and osteoporosis. Here, a biological or clinical sample suspected of containing a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide is used. However, these embodiments also

have applications to non-clinical samples, such as in the titering of antigen or antibody samples, in the selection of hybridomas, and the like.

In the clinical diagnosis or monitoring of patients with stone formation or osteoporosis, the detection of an AH-genetically-associated protein mutant, or an alteration in the levels of AH-genetically-associated protein, in comparison to the levels in a corresponding biological sample from a normal subject may be indicative of a patient with a genetic basis for AH.

However, as is known to those of skill in the art, such a clinical diagnosis would not necessarily be made on the basis of this method in isolation. Those of skill in the art are very familiar with differentiating between significant differences in types or amounts of biomarkers, which represent a positive identification, and low level or background changes of biomarkers. Indeed, background expression levels are often used to form a "cut-off" above which increased detection will be scored as significant or positive.

1. ELISAs

As detailed above, immunoassays, in their most simple and direct sense, are binding assays. Certain preferred immunoassays are the various types of enzyme linked immunosorbent assays (ELISAs) and radioimmunoassays (RIA) known in the art. Immunohistochemical detection using tissue sections is also particularly useful. However, it will be readily appreciated that detection is not limited to such techniques, and Western blotting, dot blotting, FACS analyses, and the like may also be used.

In one exemplary ELISA, the anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibodies of the invention are immobilized onto a selected surface exhibiting protein affinity, such as a well in a polystyrene microtiter plate. Then, a test composition suspected of containing the wild-type, polymorphic or mutant AH-genetically-associated protein antigen, such as a clinical sample, is added to the wells. After binding and washing to remove non-specifically bound immune complexes, the

bound wild-type, polymorphic or mutant AH-genetically-associated protein antigen may be detected. Detection is generally achieved by the addition of another anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibody that is linked to a detectable label. This type of ELISA is a simple "sandwich ELISA". Detection may also be achieved by the addition of a second anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibody, followed by the addition of a third antibody that has binding affinity for the second antibody, with the third antibody being linked to a detectable label.

In another exemplary ELISA, the samples suspected of containing the wild-type, polymorphic or mutant AH-genetically-associated protein antigen are immobilized onto the well surface and then contacted with the anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibodies of the invention. After binding and washing to remove non-specifically bound immune complexes, the bound anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibodies are detected. Where the initial anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibodies are linked to a detectable label, the immune complexes may be detected directly. Again, the immune complexes may be detected using a second antibody that has binding affinity for the first anti-wild-type, polymorphic or mutant AH-genetically-associated protein antibody, with the second antibody being linked to a detectable label.

Another ELISA in which the wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides are immobilized, involves the use of antibody competition in the detection. In this ELISA, labeled antibodies against wild-type, polymorphic or mutant AH-genetically-associated protein are added to the wells, allowed to bind, and detected by means of their label. The amount of wild-type, polymorphic or mutant AH-genetically-associated protein antigen in an unknown sample is then determined by mixing the sample with the labeled antibodies against wild-type, polymorphic or mutant AH-genetically-associated protein before or during incubation with coated wells. The presence of wild-type, polymorphic or mutant AH-genetically-associated protein in the

sample acts to reduce the amount of antibody against wild-type, polymorphic or mutant AH-genetically-associated protein available for binding to the well and thus reduces the ultimate signal. This is also appropriate for detecting antibodies against wild-type, polymorphic or mutant AH-genetically-associated protein in an unknown sample, where the unlabeled antibodies bind to the antigen-coated wells and also reduces the amount of antigen available to bind the labeled antibodies.

Irrespective of the format employed, ELISAs have certain features in common, such as coating, incubating or binding, washing to remove non-specifically bound species, and detecting the bound immune complexes. These are described as follows:

In coating a plate with either antigen or antibody, one will generally incubate the wells of the plate with a solution of the antigen or antibody, either overnight or for a specified period of hours. The wells of the plate will then be washed to remove incompletely adsorbed material. Any remaining available surfaces of the wells are then "coated" with a nonspecific protein that is antigenically neutral with regard to the test antisera. These include bovine serum albumin (BSA), casein and solutions of milk powder. The coating allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by nonspecific binding of antisera onto the surface.

In ELISAs, it is probably more customary to use a secondary or tertiary detection means rather than a direct procedure. Thus, after binding of a protein or antibody to the well, coating with a non-reactive material to reduce background, and washing to remove unbound material, the immobilizing surface is contacted with the biological sample to be tested under conditions effective to allow immune complex (antigen/antibody) formation. Detection of the immune complex then requires a labeled secondary binding ligand or antibody, or a secondary binding ligand or antibody in conjunction with a labeled tertiary antibody or third binding ligand.

"Under conditions effective to allow immune complex (antigen/antibody) formation" means that the conditions preferably include diluting the antigens and antibodies with solutions such as BSA, bovine gamma globulin (BGG) and phosphate buffered saline (PBS)/Tween. These added agents also tend to assist in the reduction of nonspecific background.

The "suitable" conditions also mean that the incubation is at a temperature and for a period of time sufficient to allow effective binding. Incubation steps are typically from about 1 to 2 to 4 hours, at temperatures preferably on the order of 25°C to 27°C, or may be overnight at about 4°C or so.

Following all incubation steps in an ELISA, the contacted surface is washed so as to remove non-complexed material. A preferred washing procedure includes washing with a solution such as PBS/Tween, or borate buffer. Following the formation of specific immune complexes between the test sample and the originally bound material, and subsequent washing, the occurrence of even minute amounts of immune complexes may be determined.

To provide a detecting means, the second or third antibody will have an associated label to allow detection. Preferably, this will be an enzyme that will generate color development upon incubating with an appropriate chromogenic substrate. Thus, for example, one will desire to contact and incubate the first or second immune complex with a urease, glucose oxidase, alkaline phosphatase or hydrogen peroxidase-conjugated antibody for a period of time and under conditions that favor the development of further immune complex formation (*e.g.*, incubation for 2 hours at room temperature in a PBS-containing solution such as PBS-Tween).

After incubation with the labeled antibody, and subsequent to washing to remove unbound material, the amount of label is quantified, *e.g.*, by incubation with a chromogenic substrate such as urea and bromocresol purple or 2,2'-azino-di-(3-ethyl-

benzthiazoline-6-sulfonic acid [ABTS] and H_2O_2, in the case of peroxidase as the enzyme label. Quantification is then achieved by measuring the degree of color generation, *e.g.*, using a visible spectra spectrophotometer.

2. Immunohistochemistry

The antibodies of the present invention may also be used in conjunction with both fresh-frozen and formalin-fixed, paraffin-embedded tissue blocks prepared for study by immunohistochemistry (IHC). For example, each tissue block consists of 50 mg of residual "pulverized" diabetic tissue. The method of preparing tissue blocks from these particulate specimens has been successfully used in previous IHC studies of various prognostic factors, and is well known to those of skill in the art (Brown *et al.*, 1990; Abbondanzo *et al.*, 1990; Allred *et al.*, 1990).

Briefly, frozen-sections may be prepared by rehydrating 50 ng of frozen "pulverized" diabetic tissue at room temperature in phosphate buffered saline (PBS) in small plastic capsules; pelleting the particles by centrifugation; resuspending them in a viscous embedding medium (OCT); inverting the capsule and pelleting again by centrifugation; snap-freezing in -70°C isopentane; cutting the plastic capsule and removing the frozen cylinder of tissue; securing the tissue cylinder on a cryostat microtome chuck; and cutting 25-50 serial sections.

Permanent-sections may be prepared by a similar method involving rehydration of the 50 mg sample in a plastic microfuge tube; pelleting; resuspending in 10% formalin for 4 hours fixation; washing/pelleting; resuspending in warm 2.5% agar; pelleting; cooling in ice water to harden the agar; removing the tissue/agar block from the tube; infiltrating and embedding the block in paraffin; and cutting up to 50 serial permanent sections.

E. Immunodetection Kits

In still further embodiments, the present invention concerns immunodetection kits for use with the immunodetection methods described above. As the wild-type,

polymorphic or mutant AH-genetically-associated protein antibodies are generally used to detect wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides, the antibodies will preferably be included in the kit. However, kits including both such components may be provided. The immunodetection kits will thus comprise, in suitable container means, a first antibody that binds to a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide, and optionally, an immunodetection reagent and further optionally, a wild-type, polymorphic or mutant AH-genetically-associated protein or peptide.

In preferred embodiments, monoclonal antibodies will be used. In certain embodiments, the first antibody that binds to the wild-type, polymorphic or mutant AH-genetically-associated protein or peptide may be pre-bound to a solid support, such as a column matrix or well of a microtitre plate.

The immunodetection reagents of the kit may take any one of a variety of forms, including those detectable labels that are associated with or linked to the given antibody. Detectable labels that are associated with or attached to a secondary binding ligand are also contemplated. Exemplary secondary ligands are those secondary antibodies that have binding affinity for the first antibody.

Further suitable immunodetection reagents for use in the present kits include the two-component reagent that comprises a secondary antibody that has binding affinity for the first antibody, along with a third antibody that has binding affinity for the second antibody, the third antibody being linked to a detectable label. As noted above, a number of exemplary labels are known in the art and all such labels may be employed in connection with the present invention.

The kits may further comprise a suitably aliquoted composition of the wild-type, polymorphic or mutant AH-genetically-associated protein or polypeptide, whether labeled or unlabeled, as may be used to prepare a standard curve for a detection assay.

The kits may contain antibody-label conjugates either in fully conjugated form, in the form of intermediates, or as separate moieties to be conjugated by the user of the kit. The components of the kits may be packaged either in aqueous media or in lyophilized form.

The container means of the kits will generally include at least one vial, test tube, flask, bottle, syringe or other container means, into which the antibody may be placed, and preferably, suitably aliquoted. Where wild-type, polymorphic or mutant AH-genetically-associated protein or a second or third binding ligand or additional component is provided, the kit will also generally contain a second, third or other additional container into which this ligand or component may be placed. The kits of the present invention will also typically include a means for containing the antibody, antigen, and any other reagent containers in close confinement for commercial sale. Such containers may include injection or blow-molded plastic containers into which the desired vials are retained.

IV. Biological Functional Equivalents

As modifications and changes may be made in the structure of wild-type, polymorphic or mutant AH-genetically-associated genes and proteins of the present invention, and still obtain molecules having like or otherwise desirable characteristics, such biologically functional equivalents are also encompassed within the present invention.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies, binding sites on substrate molecules or receptors, DNA binding sites, or such like. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence (or, of course, its underlying DNA coding sequence) and nevertheless obtain a protein with like

(agonistic) properties. It is thus contemplated by the inventors that various changes may be made in the sequence of wild-type, polymorphic or mutant AH-genetically-associated proteins or peptides, or underlying DNA, without appreciable loss of their biological utility or activity.

Equally, the same considerations may be employed to create a protein or peptide with counterveiling, *e.g.*, antagonistic properties. This is relevant to the present invention in which AH-genetically-associated mutants or analogues may be generated. For example, a AH-genetically-associated mutant may be generated and tested functionally to identify those residues important for its activity. AH-genetically-associated mutants may also be synthesized to reflect a AH-genetically-associated mutant that occurs in the human population and that is linked to the development of hypercalciuria and osteoporosis. Such mutant proteins are particularly contemplated for use in generating mutant-specific antibodies and such mutant DNA segments may be used as mutant-specific probes and primers.

In terms of functional equivalents, it is well understood by the skilled artisan that, inherent in the definition of a "biologically functional equivalent protein or peptide or gene", is the concept that there is a limit to the number of changes that may be made within a defined portion of the molecule and still result in a molecule with an acceptable level of equivalent biological activity. Biologically functional equivalent peptides are thus defined herein as those peptides in which certain, not most or all, of the amino acids may be substituted.

In particular, where shorter length peptides, it is contemplated that fewer amino acids should be made within the given peptide. Longer domains may have an intermediate number of changes. The full length protein will have the most tolerance for a larger number of changes. Of course, a plurality of distinct proteins/peptides with different substitutions may easily be made and used in accordance with the invention.

It is also well understood that where certain residues are shown to be particularly important to the biological or structural properties of a protein or peptide, *e.g.*, residues in binding regions or active sites, such residues may not generally be exchanged. Changes in domains should be carefully considered and subsequently tested to ensure maintenance of biological function, where maintenance of biological function is desired. In this manner, functional equivalents are defined herein as those peptides which maintain a substantial amount of their native biological activity.

Amino acid substitutions are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. An analysis of the size, shape and type of the amino acid side-chain substituents reveals that arginine, lysine and histidine are all positively charged residues; that alanine, glycine and serine are all a similar size; and that phenylalanine, tryptophan and tyrosine all have a generally similar shape. Therefore, based upon these considerations, arginine, lysine and histidine; alanine, glycine and serine; and phenylalanine, tryptophan and tyrosine; are defined herein as biologically functional equivalents.

To effect more quantitative changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics, these are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte & Doolittle, 1982, incorporated herein by reference). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar

biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biological functional equivalent protein or peptide thereby created is intended for use in immunological embodiments, as in certain embodiments of the present invention. U.S. Patent 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.* with a biological property of the protein.

As detailed in U.S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 ± 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).

In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

While discussion has focused on functionally equivalent polypeptides arising from amino acid changes, it will be appreciated that these changes may be effected by alteration of the encoding DNA; taking into consideration also that the genetic code is degenerate and that two or more codons may code for the same amino acid. A table of

amino acids and their codons is presented herein for use in such embodiments, as well as for other uses, such as in the design of probes and primers and the like.

In addition to the wild-type, polymorphic or mutant AH-genetically-associated peptidyl compounds described herein, the inventors also contemplate that other sterically similar compounds may be formulated to mimic the key portions of the peptide structure. Such compounds, which may be termed peptidomimetics, may be used in the same manner as the peptides of the invention and hence are also functional equivalents.

Certain mimetics that mimic elements of protein secondary structure are described in Johnson *et al.* (1993). The underlying rationale behind the use of peptide mimetics is that the peptide backbone of proteins exists chiefly to orientate amino acid side chains in such a way as to facilitate molecular interactions, such as those of antibody and antigen. A peptide mimetic is thus designed to permit molecular interactions similar to the natural molecule.

Some successful applications of the peptide mimetic concept have focused on mimetics of β-turns within proteins, which are known to be highly antigenic. Likely β-turn structure within a polypeptide can be predicted by computer-based algorithms, as discussed herein. Once the component amino acids of the turn are determined, mimetics can be constructed to achieve a similar spatial orientation of the essential elements of the amino acid side chains.

The generation of further structural equivalents or mimetics may be achieved by the techniques of modeling and chemical design known to those of skill in the art. The art of receptor modeling is now well known, and by such methods a chemical that binds to wild-type, polymorphic or mutant AH-genetically-associated protein or to a wild-type, polymorphic or mutant AH-genetically-associated protein complex can be designed and then synthesized. It will be understood that all such sterically designed constructs fall within the scope of the present invention.

V. Purification and Assays of Proteins That Interact with Either AH loci or the AH-Gentically-Associated Proteins

Certain aspects of this invention concern methods for conveniently evaluating candidate substances to identify compounds capable of interacting with a wild-type, polymorphic or mutant AH-genetically-associated protein, or even transcription of a wild-type, polymorphic or mutant AH-genetically-associated protein.

Successful candidate substances may function in the absence of mutations in AH-genetically-associated protein, in which case the candidate compound may be termed a "positive stimulator" of AH-genetically-associated protein. Alternatively, such compounds may stimulate transcription in the presence of mutated AH-genetically-associated protein, overcoming the effects of the mutation, *i.e.*, function to oppose AH-genetically-associated protein-mutant mediated AH, and thus may be termed an" AH-genetically-associated protein mutant agonist". Compounds may even be discovered which combine both of these actions. Compounds of any such class will likely be useful therapeutic agents for use in treating AH or osteoporosis.

AH-genetically-associated proteins may function by binding to DNA. One method by which to identify a candidate substance capable of stimulating AH-genetically-associated protein is based upon specific protein:protein binding. Accordingly, to conduct such an assay, one may prepare a protein with a domain and determine the ability of a candidate substance to increase the activity of the AH-genetically-associated protein to bind DNA.

Another method by which to identify a candidate substance capable of stimulating AH-genetically-associated proteins is based upon specific protein:DNA binding. Accordingly, to conduct such an assay, one would prepare an AH-genetically-associated protein and determine the ability of a candidate substance to increase the binding to a

specific DNA segment, *i.e.*, to increase the amount or the binding affinity of a specific protein:DNA complex.

Binding assays can be parallel assays, one of which contains the binding components alone and one of which contains the added candidate substance composition. One would perform each assay under conditions, and for a period of time, effective to allow the formation of protein:protein complexes or protein:DNA complexes, and one would then separate the bound complexes from any unbound protein and/or DNA and measure the amount of the complexes. An increase in the amount of any bound complex formed in the presence of the candidate substance would be indicative of a candidate substance capable of promoting AH-genetically-associated protein binding to DNA.

In such binding assays, the amount of the bound complex may be measured, after the removal of unbound species, by detecting a label, such as a radioactive or enzymatic label, which has been incorporated into the original wild-type, polymorphic or mutant AH-genetically-associated protein or even in a DNA segment. Alternatively, one could detect the protein portion of the complex by means of an antibody directed against the protein, such as those disclosed herein.

Preferred binding assays are those in which AH-genetically-associated protein is bound to a solid support and contacted with the another component to allow complex formation. Unbound protein components are then separated from the bound complexes by washing and the amount of the remaining bound complex is quantitated by detecting the label or with antibodies. Such binding assays form the basis of filter-binding and microtiter plate-type assays and can be performed in a semi-automated manner to enable analysis of a large number of candidate substances in a short period of time. Electrophoretic methods of DNA binding, such as gel-shift assays, could also be employed to separate unbound protein or DNA from bound protein:DNA complexes.

Virtually any candidate substance may be analyzed by these methods, including compounds which may interact with wild-type, polymorphic, mutant AH-genetically-associated protein, and also substances such as enzymes which may act by physically altering one of the structures present. Of course, any compound isolated from natural sources such as plants, animals or even marine, forest or soil samples, may be assayed, as may any synthetic chemical or recombinant protein.

Another potential method for stimulating AH-genetically-associated activity is to prepare a wild-type, polymorphic, mutant AH-genetically-associated protein composition and to modify the protein composition in a manner effective to increase its binding activity. The binding assays would be performed in parallel, similar to those described above, allowing the native and modified wild-type, polymorphic, mutant AH-genetically-associated protein binding to be compared. In addition to site specific mutagenesis, phosphatase and kinase enzymes may be tested, as may other agents, including proteases and chemical agents, could be employed to modify the binding properties of wild-type, polymorphic, mutant AH-genetically-associated proteins.

Cellular assays also are available for screening candidate substances to identify those capable of stimulating wild-type, polymorphic, mutant AH-genetically-associated protein and/or capable of stimulating the transcription and gene expression of AH-genetically-associated genes. A reporter gene under the control of the transcriptional regulating region of an AH-genetically-associated gene can be used. A reporter gene is a gene that confers on its recombinant host cell a readily detectable phenotype that emerges only under specific conditions.

Reporter genes are genes which encode a polypeptide not otherwise produced by the host cell which is detectable by analysis of the cell culture, *e.g.*, by fluorometric, radioisotopic or spectrophotometric analysis of the cell culture. Exemplary enzymes include luciferases, transferases, esterases, phosphatases, proteases (tissue plasminogen activator or urokinase), and other enzymes capable of being detected by their physical

presence or functional activity. A reporter gene often used is chloramphenicol acetyltransferase (CAT) which may be employed with a radiolabeled substrate, or luciferase, which is measured fluorometrically.

Another class of reporter genes which confer detectable characteristics on a host cell are those which encode polypeptides, generally enzymes, which render their transformants resistant against toxins, *e.g.*, the *neo* gene which protects host cells against toxic levels of the antibiotic G418, and genes encoding dihydrofolate reductase, which confers resistance to methotrexate. Other genes of potential for use in screening assays are those capable of transforming hosts to express unique cell surface antigens, *e.g.*, viral *env* proteins such as HIV gp120 or herpes gD, which are readily detectable by immunoassays.

The transcriptional promotion process which, in its entirety, leads to enhanced transcription is termed "activation." The mechanism by which a successful candidate substance acts is not material since the objective is to promote wild-type, polymorphic, mutant AH-genetically-associated gene expression.

To create an appropriate vector or plasmid for use in such assays one would ligate the AH-genetically-associated protein promoter and any necessary response elements to a DNA segment encoding the reporter gene by conventional methods. The relevant promoter sequences may be obtained by *in vitro* synthesis or recovered from genomic DNA and should be ligated upstream of the start codon of the reporter gene. An AT-rich TATA box region should also be employed and should be located between the sequence and the reporter gene start codon. The region 3' to the coding sequence for the reporter gene will ideally contain a transcription termination and polyadenylation site. The promoter and reporter gene may be inserted into a replicable vector and transfected into a cloning host such as *E. coli*, the host cultured and the replicated vector recovered in order to prepare sufficient quantities of the construction for later transfection into a suitable eukaryotic host.

Host cells for use in the screening assays of the present invention will generally be mammalian cells, and are preferably cell lines which may be used in connection with transient transfection studies. Cell lines should be relatively easy to grow in large scale culture. Also, they should contain as little native background as possible considering the nature of the reporter polypeptide. Examples include the Hep G2, VERO, HeLa, human embryonic kidney, 293, CHO, W138, BHK, COS-7, and MDCK cell lines, with monkey CV-1 cells being particularly preferred.

The screening assay typically is conducted by growing recombinant host cells in the presence and absence of candidate substances and determining the amount or the activity of the reporter gene. Cells containing varying proportions of candidate substances would then be evaluated for signal activation in comparison to the suppressed levels. Candidates that demonstrate dose related enhancement of reporter gene transcription or expression are then selected for further evaluation as clinical therapeutic agents.

VI. Diagnostics

As with the therapeutic methods of the present invention, the diagnostic methods are based upon the weight of evidence of the importance of AH-genetically-associated genes and other genes identified.

The diagnostic methods of the present invention generally involve detecting the presence of a particular marker or gene genetically associated with AH from a blood or urine sample from a patient with heightened susceptibility either to develop nephrolithiasis/urolithiasis or to develop osteoporosis. Once more information is known about the molecular mechanism of such markers or genes, a diagnostic method may involve determining either the type or the amount of a wild-type, polymorphic or mutant AH-genetically-associated protein present within a biological sample from a patient suspected of having AH or osteoporosis with hypercalciuria. Irrespective of the actual role of AH-genetically-associated proteins, it will be understood that the detection of

either the genetic basis for AH or a mutant protein encoded by an AH gene on human chromosome 1 is likely to be diagnostic of an increased risk of AH and osteoporosis with hypercalciuria and that the detection of altered amounts of AH-genetically-associated proteins, either at the mRNA or protein level, is also likely to have diagnostic implications, particularly where there is a reasonably significant difference in amounts.

The type or amount of a wild-type or mutant AH-genetically-associated protein present within a biological sample, such as a blood, urine, or tissue sample, may be determined by means of a molecular biological assay to determine the level of a nucleic acid that encodes such a AH-genetically-associated protein, or by means of an immunoassay to determine the level of the polypeptide itself.

Any of the foregoing nucleic acid detection methods or immunodetection methods may be employed as a diagnostic methods in the context of the present invention.

VII. Therapeutics

AH leads to elevated urinary calcium excretion. The elevated calcium salts in the urine leads to precipitate formation and eventual stone development. The underlying mechanism pertaining to the elevated calcium absorption is not understood as is the relationship of AH to osteoporosis. Osteoporosis is defined as a group of disorders that is characterized by aberrant bone remodeling; the net rate of bone resorption is greater, rather than in dynamic equilibrium with, the rate of bone formation. The condition can occur as either a primary disorder or as a disorder associated with a various disease, such as hypercalciuria. Examples of osteoporosis with hypercalciuria include ideopathic osteoporosis with hypercalciuria and postmenopausal osteoporosis with hypercalciuria. Ideopathic ostoeporosis is often times seen in young women or men demonstrating increased calcium absorption for unknown reasons. Postmenopausal osteoporosis is seen in postmenopausal women and is associated with decreased estrogen levels and increased calcium absorption.

In some families, as described later in significant detail, the risk of developing AH is associated with the presence of particular loci located on human chromosome 1. In addition to dietary adjustments and hormone therapy, more permanent therapeutic methods are also contemplated by the present invention. Depending on the underlying molecular mechanism of AH, certain therapies can be implemented to correct the AH defect.

1. Gene Therapy

If a mutant form of an AH gene is involved in the susceptibility of a patient of developing AH, a general approach of the present invention is to provide a cell with a wild-type or polymorphic AH-genetically-associated protein, thereby permitting the proper regulatory activity of the proteins to take effect. While it is conceivable that the protein may be delivered directly, a preferred embodiment involves providing a nucleic acid encoding a AH-genetically-associated protein to the cell. Following this provision, the polypeptide is synthesized by the transcriptional and translational machinery of the cell, as well as any that may be provided by the expression construct. In providing antisense, ribozymes and other inhibitors, the preferred mode is also to provide a nucleic acid encoding the construct to the cell. All such approaches are herein encompassed within the term "gene therapy".

In various embodiments of the invention, DNA is delivered to a cell as an expression construct. Several non-viral methods for the transfer of expression constructs into cultured mammalian cells also are contemplated by the present invention. These include calcium phosphate precipitation, DEAE-dextran, electroporation, direct microinjection, DNA-loaded liposomes and lipofectamine-DNA complexes, cell sonication, gene bombardment using high velocity microprojectiles, and receptor-mediated transfection. Some of these techniques may be successfully adapted for *in vivo* or *ex vivo* use, as discussed below.

In another embodiment of the invention, the expression construct may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be

performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro*, but it may be applied to *in vivo* use as well.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them. Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force. The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

In a further embodiment of the invention, the expression construct may be entrapped in a liposome, as discussed below. Also contemplated are lipofectamine-DNA complexes. Liposome-mediated nucleic acid delivery and expression of foreign DNA in vitro has been very successful. Wong *et al.* (1980) demonstrated the feasibility of liposome-mediated delivery and expression of foreign DNA in cultured chick embryo, HeLa and hepatoma cells. In certain embodiments of the invention, the liposome may be complexed with a hemagglutinating virus (HVJ). This has been shown to facilitate fusion with the cell membrane and promote cell entry of liposome-encapsulated DNA. In other embodiments, the liposome may be complexed or employed in conjunction with nuclear non-histone chromosomal proteins (HMG-1). In yet further embodiments, the liposome may be complexed or employed in conjunction with both HVJ and HMG-1. In other embodiments, the delivery vehicle may comprise a ligand and a liposome. Where a bacterial promoter is employed in the DNA construct, it also will be desirable to include within the liposome an appropriate bacterial polymerase.

The ability of certain viruses to enter cells via receptor-mediated endocytosis and to integrate into host cell genome and express viral genes stably and efficiently have

made them attractive candidates for the transfer of foreign genes into mammalian cells. Preferred gene therapy vectors of the present invention will generally be viral vectors.

Retroviruses have promise as gene delivery vectors due to their ability to integrate their genes into the host genome, transferring a large amount of foreign genetic material, infecting a broad spectrum of species and cell types and of being packaged in special cell-lines (Miller, 1992).

Other viruses, such as adenovirus, herpes simplex viruses (HSV), cytomegalovirus (CMV), and adeno-associated virus (AAV), such as those described by U.S. Patent 5,139,941, incorporated herein by reference, may also be engineered to serve as vectors for gene transfer. Although some viruses that can accept foreign genetic material are limited in the number of nucleotides they can accommodate and in the range of cells they infect, these viruses have been demonstrated to successfully effect gene expression. However, adenoviruses do not integrate their genetic material into the host genome and therefore do not require host replication for gene expression, making them ideally suited for rapid, efficient, heterologous gene expression. Techniques for preparing replication-defective infective viruses are well known in the art.

In certain further embodiments, the gene therapy vector will be HSV. A factor that makes HSV an attractive vector is the size and organization of the genome. Because HSV is large, incorporation of multiple genes or expression cassettes is less problematic than in other smaller viral systems. In addition, the availability of different viral control sequences with varying performance (temporal, strength, *etc.*) makes it possible to control expression to a greater extent than in other systems. It also is an advantage that the virus has relatively few spliced messages, further easing genetic manipulations. HSV also is relatively easy to manipulate and can be grown to high titers. Thus, delivery is less of a problem, both in terms of volumes needed to attain sufficient MOI and in a lessened need for repeat dosings.

Of course, in using viral delivery systems, one will desire to purify the virion sufficiently to render it essentially free of undesirable contaminants, such as defective interfering viral particles or endotoxins and other pyrogens such that it will not cause any untoward reactions in the cell, animal or individual receiving the vector construct. A preferred means of purifying the vector involves the use of buoyant density gradients, such as cesium chloride gradient centrifugation.

Gene delivery using second generation retroviral vectors has been reported. Kasahara *et al.* (1994) prepared an engineered variant of the Moloney murine leukemia virus, that normally infects only mouse cells, and modified an envelope protein so that the virus specifically bound to, and infected, human cells bearing the erythropoietin (EPO) receptor. This was achieved by inserting a portion of the EPO sequence into an envelope protein to create a chimeric protein with a new binding specificity.

2. Antisense

In an alternative embodiment, the AH-genetically-associated protein nucleic acids employed may actually encode antisense constructs that hybridize, under intracellular conditions, to AH-genetically-associated protein nucleic acids. The term "antisense construct" is intended to refer to nucleic acids, preferably oligonucleotides, that are complementary to the base sequences of a target DNA or RNA. Antisense oligonucleotides, when introduced into a target cell, specifically bind to their target nucleic acid and interfere with transcription, RNA processing, transport, translation and/or stability.

Antisense constructs may be designed to bind to the promoter and other control regions, exons, introns or even exon-intron boundaries of a gene. Antisense RNA constructs, or DNA encoding such antisense RNA's, may be employed to inhibit gene transcription or translation or both within a host cell, either *in vitro* or *in vivo*, such as within a host animal, including a human subject. Nucleic acid sequences which comprise "complementary nucleotides" are those which are capable of base-pairing according to

the standard Watson-Crick complementarity rules. That is, that the larger purines will base pair with the smaller pyrimidines to form combinations of guanine paired with cytosine (G:C) and adenine paired with either thymine (A:T), in the case of DNA, or adenine paired with uracil (A:U) in the case of RNA. Inclusion of less common bases such as inosine, 5-methylcytosine, 6-methyladenine, hypoxanthine and others in hybridizing sequences does not interfere with pairing.

As used herein, the terms "complementary" means nucleic acid sequences that are substantially complementary over their entire length and have very few base mismatches. For example, nucleic acid sequences of fifteen bases in length may be termed complementary when they have a complementary nucleotide at thirteen or fourteen positions with only a single mismatch. Naturally, nucleic acid sequences which are "completely complementary" will be nucleic acid sequences which are entirely complementary throughout their entire length and have no base mismatches.

Other sequences with lower degrees of homology also are contemplated. For example, an antisense construct which has limited regions of high homology, but also contains a non-homologous region (*e.g.*, a ribozyme) could be designed. These molecules, though having less than 50% homology, would bind to target sequences under appropriate conditions.

While all or part of the AH-genetically-associated protein gene sequence may be employed in the context of antisense construction, short oligonucleotides are easier to make and increase *in vivo* accessibility. However, both binding affinity and sequence specificity of an antisense oligonucleotide to its complementary target increases with increasing length. One can readily determine whether a given antisense nucleic acid is effective at targeting of the corresponding host cell gene simply by testing the constructs *in vitro* to determine whether the function of the endogenous gene is affected or whether the expression of related genes having complementary sequences is affected.

In certain embodiments, one may wish to employ antisense constructs which include other elements, for example, those which include C-5 propyne pyrimidines. Oligonucleotides which contain C-5 propyne analogues of uridine and cytidine have been shown to bind RNA with high affinity and to be potent antisense inhibitors of gene expression.

VIII. Pharmaceutical Compositions

A. Pharmaceutically Acceptable Carriers

Aqueous compositions of the present invention comprise an effective amount of the AH-genetically-associated protein, peptide, epitopic core region, inhibitor, or such like, dissolved or dispersed in a pharmaceutically acceptable carrier or aqueous medium. Aqueous compositions of gene therapy vectors expressing any of the foregoing are also contemplated. The phrases "pharmaceutically or pharmacologically acceptable" refer to molecular entities and compositions that do not produce an adverse, allergic or other untoward reaction when administered to an animal, or a human, as appropriate.

As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

For human administration, preparations should meet sterility, pyrogenicity, general safety and purity standards as required by FDA Office of Biologics standards.

The biological material should be extensively dialyzed to remove undesired small molecular weight molecules and/or lyophilized for more ready formulation into a desired vehicle, where appropriate. The active compounds will then generally be formulated for

parenteral administration, *e.g.*, formulated for injection via the intravenous, intramuscular, subcutaneous, intralesional, or even intraperitoneal routes. The preparation of an aqueous composition that contains an AH-genetically-associated agent as an active component or ingredient will be known to those of skill in the art in light of the present disclosure. Typically, such compositions can be prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for using to prepare solutions or suspensions upon the addition of a liquid prior to injection can also be prepared; and the preparations can also be emulsified.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions; formulations including sesame oil, peanut oil or aqueous propylene glycol; and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi.

Solutions of the active compounds as free base or pharmacologically acceptable salts can be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

An AH-genetically-associated protein, peptide, agonist or antagonist of the present invention can be formulated into a composition in a neutral or salt form. Pharmaceutically acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium,

calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like.

The carrier can also be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

In terms of using peptide therapeutics as active ingredients, the technology of U.S. Patents 4,608,251; 4,601,903; 4,599,231; 4,599,230; 4,596,792; and 4,578,770, each incorporated herein by reference, may be used.

The preparation of more, or highly, concentrated solutions for direct injection is also contemplated, where the use of DMSO as solvent is envisioned to result in extremely rapid penetration, delivering high concentrations of the active agents to a small tumor area.

Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms, such as the type of injectable solutions described above, but drug release capsules and the like can also be employed.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, sterile aqueous media which can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage could be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject.

The active AH-genetically-associated protein-derived peptides or agents may be formulated within a therapeutic mixture to comprise about 0.0001 to 1.0 milligrams, or about 0.001 to 0.1 milligrams, or about 0.1 to 1.0 or even about 10 milligrams per dose or so. Multiple doses can also be administered.

In addition to the compounds formulated for parenteral administration, such as intravenous or intramuscular injection, other pharmaceutically acceptable forms include,

e.g., tablets or other solids for oral administration; liposomal formulations; time release capsules; and any other form currently used, including cremes.

One may also use nasal solutions or sprays, aerosols or inhalants in the present invention. Nasal solutions are usually aqueous solutions designed to be administered to the nasal passages in drops or sprays. Nasal solutions are prepared so that they are similar in many respects to nasal secretions, so that normal ciliary action is maintained. Thus, the aqueous nasal solutions usually are isotonic and slightly buffered to maintain a pH of 5.5 to 6.5.

In addition, antimicrobial preservatives, similar to those used in ophthalmic preparations, and appropriate drug stabilizers, if required, may be included in the formulation. Various commercial nasal preparations are known and include, for example, antibiotics and antihistamines and are used for asthma prophylaxis.

Additional formulations which are suitable for other modes of administration include vaginal suppositories and pessaries. A rectal pessary or suppository may also be used.

Suppositories are solid dosage forms of various weights and shapes, usually medicated, for insertion into the rectum, vagina or the urethra. After insertion, suppositories soften, melt or dissolve in the cavity fluids.

In general, for suppositories, traditional binders and carriers may include, for example, polyalkylene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1%-2%.

Vaginal suppositories or pessaries are usually globular or oviform and weighing about 5 g each. Vaginal medications are available in a variety of physical forms, *e.g.*, creams, gels or liquids, which depart from the classical concept of suppositories. Vaginal

tablets, however, do meet the definition, and represent convenience both of administration and manufacture.

Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders.

In certain defined embodiments, oral pharmaceutical compositions will comprise an inert diluent or assimilable edible carrier, or they may be enclosed in hard or soft shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet. For oral therapeutic administration, the active compounds may be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 0.1% of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 2 to about 75% of the weight of the unit, or preferably between 25-60%. The amount of active compounds in such therapeutically useful compositions is such that a suitable dosage will be obtained.

The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup of

elixir may contain the active compounds sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor.

B. Liposomes and Nanocapsules

In certain embodiments, the use of liposomes and/or nanoparticles is contemplated for the introduction of wild-type, polymorphic or mutant AH-genetically-associated protein peptides or agents, or gene therapy vectors, including both wild-type and antisense vectors, into host cells. The formation and use of liposomes is generally known to those of skill in the art, and is also described below.

Nanocapsules can generally entrap compounds in a stable and reproducible way. To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention, and such particles may be are easily made.

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μm. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

The following information may also be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely

packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

Liposomes interact with cells via four different mechanisms: Endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or *vice versa*, without any association of the liposome contents. Varying the liposome formulation can alter which mechanism is operative, although more than one may operate at the same time.

C. Kits

Therapeutic kits of the present invention are kits comprising a wild-type, polymorphic or mutant AH-genetically-associated protein, peptide, inhibitor, gene, vector or other AH-genetically-associated protein effector. Such kits will generally contain, in suitable container means, a pharmaceutically acceptable formulation of an AH-genetically-associated protein, peptide, domain, inhibitor, or a gene or vector expressing any of the foregoing in a pharmaceutically acceptable formulation, optionally comprising other anti-AH or osteoporosis agents. The kit may have a single container means, or it may have distinct container means for each compound.

When the components of the kit are provided in one or more liquid solutions, the liquid solution is an aqueous solution, with a sterile aqueous solution being particularly preferred. The AH-genetically-associated protein compositions may also be formulated into a syringeable composition. In which case, the container means may itself be a syringe, pipette, or other such like apparatus, from which the formulation may be applied

to an infected area of the body, injected into an animal, or even applied to and mixed with the other components of the kit.

However, the components of the kit may be provided as dried powder(s). When reagents or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent. It is envisioned that the solvent may also be provided in another container means.

The container means will generally include at least one vial, test tube, flask, bottle, syringe or other container means, into which the AH-genetically-associated protein or gene or inhibitory formulation are placed, preferably, suitably allocated. Where a second AH therapeutic is provided, the kit will also generally contain a second vial or other container into which this agent may be placed. The kits may also comprise a second/third container means for containing a sterile, pharmaceutically acceptable buffer or other diluent.

The kits of the present invention will also typically include a means for containing the vials in close confinement for commercial sale, such as, *e.g.*, injection or blow-molded plastic containers into which the desired vials are retained.

Irrespective of the number or type of containers, the kits of the invention may also comprise, or be packaged with, an instrument for assisting with the injection/administration or placement of the ultimate AH-genetically-associated protein or gene composition within the body of an animal. Such an instrument may be a syringe, pipette, forceps, or any such medically approved delivery vehicle.

IX. Examples

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor

to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Example 1

Clinical Evaluation of Absorptive Hypercalciuria (AH)

This example demonstrates the clinical evaluation of three kindreds with severe AH. Examination of multiple clinical parameters was used for phenotypic assignment.

METHODS

Kindred Description and Evaluation

All participants gave informed consent to a protocol approved by the Institutional Review Board. Three kindreds with severe AH participated in this study. The probands were identified from patients in the kidney stone clinic. In the first kindred, AH-01 (FIG. 1A), 22 family members and 4 unrelated spouses were evaluated. In the second kindred, AH-02 (FIG. 1B), 5 individuals were evaluated and in the third kindred, AH-03 (FIG. 1C), 4 family members were evaluated. All kindreds were North American Caucasians of Western European descent. The number of subjects evaluated and the scope of investigation depended on the willingness and cooperation of the subjects. Either an inpatient or outpatient evaluation was done on consenting study participants. Some individuals agreed to undergo only a partial outpatient evaluation.

Inpatient Evaluation

Patients were evaluated according to established protocols (Breslau *et al.*, 1992; Pak *et al.*, 1974). Individuals were admitted to the General Clinical Research Center for 4 days where they were maintained on a constant metabolic diet containing 100 mmol sodium, 10 mmol calcium and 25.8 mmol phosphorous per day for 3 days (Days 1-3) after being on an instructed diet of similar composition for 1 week prior to admission. Fasting blood samples on days 1 - 4 were analyzed for calcium and alkaline phosphatase

(Smith-Kline Beecham, Dallas, TX). Fasting venous blood samples on days 1 and 4 were analyzed for serum iPTH by immunoradiometric assay (Nichols Institute, San Juan Capistrano, CA) and $1,25(OH)_2D$ (by radioreceptor assay). Calcium and creatinine were measured in three successive 24-h urine collections (Days 1 to 3). On Day 4, a 2 h fasting urine collection was obtained for measurement of calcium and creatinine, and a 4 h urine collection was obtained for the same tests after an oral ingestion of a synthetic meal containing 1 g of calcium (Pak *et al.*, 1980; Pak *et al.*, 1975). The calciuric response after the calcium load gave an indirect measure of intestinal calcium absorption (Pak *et al.*, 1980; Pak *et al.*, 1975). Fractional calcium absorption (α) was determined either from the fecal recovery of ^{47}Ca after ingestion of a synthetic test meal containing trace radiocalcium, (Pak *et al.*, 1974) or by using a double stable isotope technique (Abrams, 1994). The two tests yielded equivalent results. Bone mineral density of L2-L4 vertebrae, femoral neck and radial shaft was measured using dual energy x-ray absorptiometry (Hologic QDR-2000, Waltham MA). A heparinized venous blood was obtained for lymphocyte isolation and immortalization and an EDTA treated venous blood for genomic DNA isolation (Neitzel, 1989).

Outpatient Evaluation

Subjects underwent an outpatient evaluation (Pak *et al.*, 1989) following one week on an instructed diet designed to mimic the inpatient metabolic diet in sodium, calcium and phosphorous content. This evaluation included: fasting venous serum for calcium, creatinine, iPTH and $1,25(OH)_2D$, heparinized venous blood for lymphocyte isolation and immortalization, EDTA treated venous blood for genomic DNA isolation, a 24-h urine collection for calcium and creatinine, a 2 h fasting urine collection was obtained for measurement of calcium and creatinine, and a 4 h urine collection was obtained for the same tests after an oral ingestion of a synthetic meal containing 1 g of calcium (Pak *et al.*, 1980; Pak *et al.*, 1975). Each participant completed a standardized questionnaire that included kidney stone and dietary history.

Phenotype Assignment

Phenotype assignment in kindreds AH-01 and AH-02 was based on 4 criteria: (1) evidence of hyperabsorption of calcium, either a calciuric response to an oral calcium load > 0.20 mg Ca/dl glomerular filtrate (GF) or α > 61%, (2) elevated fasting urinary calcium (> 0.11 mg Ca /dl GF), (3) hypercalciuria (> 200 mg Ca/ day on a calcium restricted diet) and (4) a low or normal serum PTH (< 65 pg/ml) (Levy *et al.*, 1995). Individuals who satisfied at least three of the four criteria were assigned affected phenotype. Those with intestinal hyperabsorption of calcium (criterion 1) who met only one additional criterion were classified as unknown phenotype. If an unrelated spouse had either an AH phenotype or was not evaluated, their progeny, who would otherwise have an affected or unknown phenotype, were assigned unknown phenotype. All others were classified as unaffected.

In kindred AH-03, affected phenotype assignment was based on the satisfaction of criteria 3 and 4 alone, since fasting urinary calcium, calciuric response to an oral calcium load and were done only on the proband. An unknown status was assigned when only criterion 3 was met.

All affected members from all 3 kindreds had normocalcemia. Bone density was not utilized in the definition of AH phenotype since only a limited number of subjects were available for this measurement.

RESULTS

Probands

The proband of the kindred AH-01 (FIG. 1A, III-14) was a 37 year old white male who underwent an outpatient evaluation. He had a history of recurrent kidney stone formation, elevated 24 h urine calcium, fasting urinary calcium and calcium load response and a low serum iPTH. The proband of the kindred AH-02 (FIG. 1B, III-2) was a 47 year old white female who underwent an inpatient evaluation. She had elevated 24 h urine calcium, fasting urinary calcium, calcium load response and • and a normal iPTH. The proband of the kindred AH-03 (FIG. 1C, III-4) was a 32 year old white male who

underwent an inpatient evaluation. He had a history of recurrent kidney stone formation, elevated 24 h urine calcium, calcium load response and α and a high normal fasting urinary calcium and a normal iPTH. All three probands had no history of bowel disease, primary hyperparathyroidism, primary hyperoxaluria, renal tubular acidosis, gout or cystinuria. They all satisfied the diagnostic criteria of AH (Levy *et al.*, 1995; Breslau *et al.*, 1992).

Families

Kindred AH-01 (FIG. 1A). Twenty-six blood samples, including the proband, were collected for genotype analysis. Twenty-four members of the family underwent clinical evaluation using the outpatient protocol. Bone density measurements were obtained on eight family members and three unrelated spouses. Biochemical and physiological characteristics of family members with affected phenotype are presented in Table 4.

The 12 affected family members including the proband had biochemical features of severe AH, (Levy *et al.*, 1980; Breslau *et al.*, 1992) with markedly elevated 24 h urinary Ca, fasting hypercalciuria, notably exaggerated calciuric response to oral Ca load, and low normal serum PTH. Serum calcium and 1,25-$(OH)_2$D were normal. Bone density was low (Table 4).

There were 7 stone-formers in the family (FIG. 1A) and 1 stone-forming spouse (III-2). Stones that were analyzed were calcium oxalate and/or calcium phosphate in composition. Six of the 7 stone-formers had the affected phenotype (FIG. 1A), while the remaining stone-former (IV-I) was of uncertain phenotype. Based on clinical evaluation, 5 additional non-stone-forming family members and one spouse had the AH phenotype (FIG. 1A). Individuals III-18 and III-19 were monozygotic twins and were treated as a single entity for purposes of linkage analysis.

Kindred AH-02 (FIG. 1B). Five individuals underwent phenotypic evaluation and genotype analysis. Three members of the kindred had inpatient evaluations with the

determination of α, while the remaining two had outpatient evaluations. Bone density measurements were obtained on three family members. The three affected family members, including the proband (FIG. 1B), had evidence of severe AH, with fasting hypercalciuria and low bone density (Table 4). One member was assigned unknown phenotype (III-1). There was a family history of stone formation on the maternal side of the family (I-3 and II-4).

Kindred AH-03 (FIG. 1C). Four members of the kindred underwent phenotypic evaluation and genotype analysis. The proband had an inpatient evaluation while the three family members had a partial outpatient evaluation that excluded the determination of fasting urinary calcium and the calciuric response to a calcium load. Bone density measurements were obtained in three family members. The three affected individuals, including the proband (II-1, III-2 and III-4, FIG. 1C), had biochemical features compatible with severe AH (Table 4). They had marked hypercalciuria, and low bone density. One member (III-5) was assigned an unknown phenotype. Of the subjects evaluated, only the proband had nephrolithiasis although a paternal cousin (III-7) also reported a history of nephrolithiasis.

TABLE 4
Mean Biochemical and Physiological Characteristics of Affected Individuals From Study Kindreds

	Family			
Parameter	**AH-01**	**AH-02**	**AH-03**	**Normal Range**
Serum				
IPTH, ng/L	15±8[a] (10)	38±10 (3)	31±2 (3)	10-65
Alkaline Phosphatase (IU/L)	61±13 (10)	62±30 (2)	76±23 (3)	30-140
1,25 $(OH)_2D$, pmol/L	115±26 (10)	82±24 (3)	84±7 (3)	48-132
Ca, mmol/L	2.3±0.05 (10)	2.1±0.1 (3)	2.4±0.03 (3)	
Urine				
24 hour Ca, nmol/day	7.3±2.2 (11)	6.1±1.0 (3)	7.0±0.9 (3)	<5.0
Fasting Ca, mg/dL GF	0.037±0.015 (10)	0.0370±0.015 (3)	0.022 (1)	<0.027
Post Ca load, mg/dL GF	0.097±0.030 (9)	0.070±0.010 (3)	0.065 (1)	<0.050
Intestinal Ca Absorption				
α, %	ND	67.9±7.7 (3)	68.8 (1)	40-60
Bone Density				
L2-L4				
BMD (g/cm^2)	1.02±0.09 (9)	0.748±0.07 (3)	0.88±0.04 (2)	
T Score	-0.82±0.81	-3.10±0.68	-2.10±0.35	
Z Score	-041±0.78	-1.90±0.91	-1.95±0.13	
Femur (Neck)				
BMD	0.83±0.14 (9)	0.53±0.03 (3)	0.84±0.11 (2)	
T Score	-0.46±1.22	-2.84±0.26	-0.69±0.82	
Z Score	-0.14±0.76	-1.79±1.23	-0.48±0.78	
Radius				
BMD	0.61±0.09 (8)	0.57±0.08 (3)	0.76±0.03 (2)	
T Score	-2.91±1.56	-1.97±1.29	-1.08±0.53	
Z Score	-2.41±1.42	-0.62±0.99	-0.79±0.52	

[a] All values are presented mean ± SD with number of samples given in ( ). ND, not determined.

Example 2

Determination of the Chromosomal Locus that is Linked to Intestinal Calcium Hyperabsorption in Three Kindreds with Severe AH.

The presence of a family history of nephrolithiasis, in about half of the affected individuals studied, indicates that an inherited genetic defect is one likely cause of AH. While it is known that intestinal calcium absorption is regulated by a number of factors, the molecular biological basis for the increased calcium absorption in AH is unknown. This example demonstrates the identification of the chromosomal locus of the gene defect linked to the AH phenotype in three families with a severe form of AH.

METHODS

DNA Analysis

Genomic DNA was prepared from peripheral blood lymphocytes (Qiagen, Valencia, CA) DNA genotyping was performed using fluorescently labeled primers available from Perkin-Elmer Applied Biosystems (ABI), or Research Genetics, on an ABI Model 377 automated DNA sequencer with GENESCAN 2.0 software. A total of 178 randomly spaced markers (10 -30 cM spacing) were analyzed in the initial low density screening. Regions where a two-point lod score was >0.3 were screened using high density markers. Fifty-five additional markers were used in this secondary screening. All PCR™ amplification reactions were performed in a Perkin Elmer thermal cycler, Model 9600 following suppliers protocols. Samples were analyzed on 4% polyacrylamide gels. Data analysis was performed using GENOTYPER software (ABI).

Linkage analysis

Two point lod scores were calculated using the computer program Linkage 5.1 (Lathrop *et al.*, 1985). The AH trait was assumed to be dominant with a penetrance of 80% and a disease frequency of 0.02. Non-parametric multipoint linkage analysis was performed using GENEHUNTER software (Whitehead Institute for Biomedical Research) (Kruglyak *et al.*, 1996). Analyses were run on a 180 MHz Pentium Pro

computer using max bits = 20. All affected individuals were included in the analysis. Allele frequencies and map distances were taken from the literature (Collins *et al.*, 1996; CEPH genotype database, Ceph-Genethon internet site http//www.cephb.fr.).

RESULTS

Linkage analysis

Kindred AH-01 was first tested for linkage at potential candidate gene loci, which included genes coding for the vitamin D receptor, 1-α hydroxylase, plasma membrane calcium ATPase (PMCa1), calbindin 28K, PTHrp, NPT1, NPT2, osteocalcin, IL-1α, IL-1α and IL-1 receptor. No evidence for linkage was found at any of these loci. Candidate genes located on the X chromosome, such as CLCN5 and calbindin 9K, were eliminated, since male-to-male transmission was present in kindreds AH-01 and AH-03 ruling out a sex-linked gene defect (FIG. 1A and FIG. 1C). After elimination of these candidate gene loci, a genome wide screen was undertaken. Strong evidence of linkage was found only on the q-arm of chromosome 1 after analyzing 178 markers randomly distributed at 10 - 30 cM intervals within the genome. An additional 55 high density markers were analyzed in regions where a lod score of > 0.3 was obtained.

The maximum two point lod scores calculated for chromosomes 2 to 22 are shown in Table 5. None exceeded 1.3. However, on chromosome 1, a positive two point lod score of 2.7 was obtained for kindred AH-01 between marker D1S196 and the AH phenotype at $\theta = 0$ (Table 6). Values are given separately for each of the individiual families, AH-01, AH-02 and AH-03, and also as the combined score for all three families. Combination of the three kindreds gave a two point lod score of 3.3 (Table 6).

TABLE 5

Maximum 2 point lod scores calculated for each chromosome between microsatellite markers and AH phenotype.

Chromosome	Marker	Z max	θ max
2	D2S396	0.8	0.2
3	D3S1565	0.8	0.2
4	D4S391	0.8	0.2
5	D5S422	0.4	0.2
6	D6S422	0.4	0.2
7	D7S515	0.1	0
8	D8S260	0.92	0.1
9	D9S161	0.01	0.4
10	D10S537	0.7	0.2
11	D11S935	-0.1	0.4
12	D12S80	0.9	0.1
13	D13S158	-0.1	0.4
14	D14S280	0.1	0.4
15	D15S131	1.3	0.1
16	D16S401	-0.1	0.4
17	D17S799	-0.2	0.4
18	D18S559	0.3	0.4
19	D19S418	0.3	0.3
20	D20S117	0.1	0.4
21	D21S263	-0.1	0.4
22	D22S280	0.02	0.4

TABLE 6

Critical 2 point lod scores between microsatellite markers and AH phenotype.

Marker	Recombination Fraction, θ						Z max	θ max
	0	0.01	0.05	0.1	0.2	0.3		
AH-01								
D1S426	-0.7	-0.7	-0.5	-0.2	0.03	0.08	0.08	0.28
D1S2681	0.4	0.4	0.5	0.6	0.6	0.5	0.7	0.17
D1S196	2.7	2.7	2.5	2.3	1.8	1.2	2.7	0
D1S2815	0.6	1.0	1.3	1.4	1.2	0.8	1.4	0.09
AH-02								
D1S426	0.3	0.3	0.2	0.2	0.1	0.06	0.3	0
D1S2681	0.3	0.3	0.2	0.2	0.1	0.1	0.3	0
D1S196	0.3	0.3	0.2	0.2	0.1	0.1	0.3	0
D1S2815	8×10^{-4}	8×10^{-4}	7×10^{-4}	5×10^{-4}	3×10^{-4}	1×10^{-4}	8×10^{-4}	0
AH-03								
D1S426	-1.1	-0.9	-0.6	-0.4	-0.2	-0.07	-0.02	0.4
D1S2681	0.3	0.3	0.3	0.2	0.1	0.1	0.3	0
D1S196	0.3	0.3	0.3	0.2	0.1	0.1	0.3	0
D1S2815	0.3	0.3	0.3	0.2	0.1	0.1	0.3	0
Combined								
D1S426	-1.2	-1.1	-0.7	-0.4	0	0.1	0.1	0.3
D1S2681	-0.3	-0.2	0.2	0.4	0.6	0.5	0.6	0.2
D1S196	3.3	3.2	3.0	2.7	2.0	1.3	3.3	0
D1S2815	0.9	1.2	1.5	1.6	1.3	0.9	1.6	0.08

Using data from all 3 kindreds, high density mapping was performed using 13 additional markers, and multipoint linkage analysis was conducted over the region spanning the D1S196 locus. Several markers were either non-informative or partially informative and were not included in the multipoint analysis. Multipoint non-parametric analysis of the data yielded a non-parametric lod (NPL) score of 12.7 (p=0.000006) between markers D1S318 and D1S431 (FIG. 2B). This locus corresponded to a region contained in 1q23.3-q24, based on current mapping location for these markers (Collins *et al.*, 1996). Since the phenotype workup of the third kindred was not as complete as for the other two kindreds, a multipoint analysis using only the first two kindreds was done. An NPL score of 15.2 (p = 0.000007) was obtained for kindreds AH-01 and AH-02.

Haplotypes were constructed and analyzed for informative recombinations using the markers D1S426, D1S2681, D1S318, D1S196, D1S431 D1S2750, D1S2799, D1S2815, D1S218, D1S416 and D1S466 (FIG. 2A). Key recombinational events in individuals III-18 and IV-7 from kindred AH-01 delineated the AH gene locus between markers D1S2681 (centromere) and D1S2815 (telomere). This locus corresponded to a physical map distance of approximately 4.3 cM based on the current Linkage Data Base (LDB) composite map (Collins *et al.*, 1996. Six additional individuals (II-2, III-3, III-15, IV-1, IV-2 and IV-8) from kindred AH-01 carried the affected genotype based on haplotype analysis. Five of these individuals, including one stone-former, had previously been classified as uncertain phenotype, while the sixth individual had not undergone a clinical evaluation. All phenotypically affected individuals carried the disease genotype.

DISCUSSION

The inventors have identified a single locus on chromosome 1q23.3-q24 linked to an AH phenotype in three unrelated kindreds with AH. Members of all three kindreds who were classified as phenotypically affected met the diagnostic criteria for AH (Pak *et al.*, 1980; Levy *et al.*, 1995. The common genotypes at the chromosome 1q23.3-q24 locus were identified in all of the related members with stones.

The clinical presentations of AH in all three kindreds were compatible with a severe form of AH. Thus, their characteristic features were a moderate to marked hypercalciuria, low bone density and fasting hypercalciuria. However, there is some evidence that the molecular abnormality disclosed here may be more generalized. Fasting hypercalciuria is not an uncommon finding in AH and may be present in a substantial number of patients, especially in those with marked intestinal hyperabsorption of calcium and parathyroid suppression (Preminger *et al.*, 1989; Breslau *et al.*, 1992; Pak and Galosy, 1979; Heller *et al.*, 1998. In addition, low spinal bone density (Bataille *et al.*, 1991; Barkin *et al.*, 1985) was present in AH patients with normal fasting urinary calcium (Pietchmann *et al.*, 1992) as well as the subgroup with fasting hypercalciuria.

Despite a rich family history of kidney stone formation in patients with AH, controversy persists concerning the mode of inheritance of this disease (Coe *et al.*, 1979; Pak *et al.*, 1981; Resnick *et al.*, 1968). The inventors therefore used a non-parametric model-independent method of analysis (Genehunter) (Kruglyak *et al.*, 1996) as well as a parametric method of analysis that assumed an autosomal dominant mode of inheritance (Coe *et al.*, 1979; Pak *et al.*, 1981). The results of the genome wide screening, using both methods of analysis, indicated that only one region of the genome met the criteria for linkage. Thus, the inventors conclude that AH is inherited in an autosomal dominant mode due to a gene mutation in the chromosome 1q23.3-q24 locus, at least in the three kindreds evaluated. Based on the most recent chromosome 1 map, no genes of known function have been identified in this candidate region. This lack of a known calcium-regulatory gene at this chromosomal locus leads to the intriguing possibility that an, as yet, unreported gene may be involved in the regulation of intestinal calcium absorption and possibly bone loss.

Prior pathogenetic mechanisms for AH have implicated an abnormality in either vitamin D metabolism or the vitamin D receptor Breslau *et al.*, 1992; Insogna *et al.*, 1985; Zerwekh *et al.*, 1993; Krieger *et al.*, 1996. However, both the vitamin D receptor and 1-α hydroxylase gene, which catalyzes the formation of $1,25(OH)_2D$, have been

eliminated as candidates in the three families based on the linkage data. Several other candidate genes, including PMCA1, PMCA4, the 28K and 9K calbindins and the Na/Pi cotransporter genes, NPT1 and NPT2 (Tieder *et al.*, 1985; Tenehouse, 1997) have been implicated in the regulation of either the cellular transport of calcium or the renal excretion of calcium. However, all of these genes were also eliminated from being involved in the etiology of AH in the families studied, since their reported chromosomal loci did not correspond to chromosome 1q23.3-q24.

Some family members without stones also had the common genotype at chromosome 1q23.3-q24. The incomplete penetrance of stone formation is likely due to the influence of environmental factors or possibly to other disease modifying genes. The inventors chose not to use stone formation as part of the phenotype in order to prevent other factors from complicating the analysis.

Identification of the specific gene and mutations contained therein will be necessary in order to determine both the relationship of this gene defect to the clinical features associated with AH and the prevalence of this gene defect in the AH patient population.

Example 3

Methods for Isolation and Characterization of a Gene Linked to AH

The initial identification of YACs (yeast artificial chromosomes) derived from the chromosomal region containing the disease gene is done by querying the CEPH/Genethon database. There are 8 YACs available spanning the D1S196 region and these have been ordered into a contiguous region ("contig"). YACs are available from Genome Systems, Inc., or Research Genetics. An example of a detailed protocol for preparation of YAC DNA in liquid form for PCR analysis or agarose plugs known to those of skill in the art is described by Horrigan and Westbrook (1997). The suspected overlaps in the YACs are confirmed by fingerprinting and endclone rescue. Human specific sequences are identified by interspersed repeated sequence PCR (IRS-PCR) (Ledbetter *et al.*, 1990).

Generation of IRS-PCR fingerprints is performed as follows: a 50 µl PCR reaction will be set up to contain 5-10 ng YAC DNA, 10X XL buffer II, 1.5 mM $Mg(OAc)_2$, 200 mM dNTPs, ng IRS primer (*e.g.* ALU3' GAT CGC GCC ACT GCA CTC C, SEQ ID NO:3, and ALU 5 GGA TTA CAG GCG TGA GCC AC, SEQ ID NO:4) total for all primers and 2.5 U rTth XL polymerase (Perkin Elmer); typical cycling conditions are 94°C for 3 minutes then 25 cycles of 94°C for 40 seconds, 60°C for 1 minute, 72°C for 5 minute a final 10 minute elongation at 72°C is performed. Products are then analyzed by separation on a 1.5% agarose gel stained with ethidium bromide (EtBr). If overlap of YACs is not apparent, the ends of the YAC clone will be isolated by IRS-vector PCR (Fujita and Swaroop, 1995). At this stage the YAC may be converted into a series of smaller overlapping cosmid clones for maintenance. DNA derived by IRS-PCR can be used directly as a hybridization probe to select clones on high density filters containing a human genomic DNA library. Briefly DNA products will be purified on Qiax spin columns and labelled by random priming. Probes will be combined and prehybridized with 50 µg/ml Cot-1 DNA and 100 mg/ml sonicated human placental DNA in 500 ml of 0.12 M Na_2HPO_4 for 4 hours at 68°C. Filters containing the target clones are prehybridized in 0.5 M Na_2HPO_4, 7% SDS, 1 mM EDTA, 50 mg/ml denatured sonicated salmon sperm DNA for 4 hours at 68°C. Filters will be hybridized with 5x 10^6 cpm/ml probe for 18-24 hours at 68°C. Filters will be washed at 68°C for 30 minutes in 40 mM Na_2HPO_4, 5% SDS, 1 mM EDTA and then twice for 30 minutes in 40 mM Na_2HPO_4, 1% SDS, 1mM EDTA. X-ray fill will be exposed to the filters for 1-2 days at -70°C with intensifying screens. Genes contained within the YACs will be identified by exon trapping; both internal and 3' exons will be isolated as described in detail by Krizman (1997). Internal exon sequences will be used for coding region searches of databases for identification of similarities with other known proteins in an attempt to identify the type of gene product. 3' terminal sequences can be used for direct screening of cDNAs. pSPL3 is a vector of choice for internal exon trapping and replication will be performed in Cos 7 cells as previously described (Krizman 1997). Once the entire sequence of a cDNA or gene has been determined, sequencing primers are prepared for analysis of the gene in patient samples of DNA. Comparison of the patient and wildtype DNA

sequences will allow identification of any mutations present in AH DNA. More rapid screening of patient cDNA or genomic DNA can be achieved by RNA mismatch cleavage analysis. Depending on the nature of the mutation, a prediction can be made regarding the effect of the mutation on the protein structure. Coincidence of mutation and phenotype will be confirmed by analyzing DNA from all affected and non-affected family members. Expression studies will be performed to confirm the effect of the mutation on protein function.

Example 4

Detection of a Mutation in the AH Loci to Identify Individuals at Risk for AH

An example of a typical technique for screening for mutations is based on PCR amplification of the identified region or regions of the AH gene containing the mutated sequence, if the mutation involves either an insertion or a deletion. For instance, a typical protocol would involve preparing genomic DNA from 1 ml whole blood sample from a patient using a Qiagen micro-DNA preparation kit (Qiagen, Valencia, CA). DNA genotyping will be performed using fluorescently-labeled primers designed to flank known region(s) containing the mutated sequence with an ABI Model 377 automated DNA sequencer with GENESCAN 2.0 software. All PCR amplification reactions will be performed in a 9600 Perkin Elmer thermal cycler. Typical conditions for such mutliplex PCR will be as follows: 60 ng of genomic DNA will be amplified by mutiplex PCR in a total volume of 15 μl, containing PCR buffer (Perkin Elmer), 2.5 mM $MgCl_2$, 0.25 units of *AmpliTaq Gold* DNA polymerase (Perkin Elmer), 250 nM dNTPs and 330 nM primers. The following cycling conditions will be used: 95°C for 10 min followed by 10 cycles (95°C for 15 s, 55°C for 15 s, 72°C for 30 s) followed by 35 cycles of (89°C for 15s, 55°C for 15 s, 72°C for 30 s) and a final elongation step of 72°C for 10 min.). Analysis of the resulting PCR products will be performed by separation on a denaturing polyacrylamide gel. Mutant DNA will be identified based on size through comparison to both wildtype and known samples of mutant DNA.

If the mutation does not involve either a deletion or insertion but is instead a base substitution, the preferred diagnostic testing will involve RNase cleavage mismatch analysis. cDNA will be analyzed if the mutation resides within the coding region. The illegitimate transcription technique (Chelly *et al.,* 1989) will be used to prepare specific mRNA for cDNA analysis. Briefly, mRNA will be prepared from 0.5-1.0 x 10^7 lymphocytes using a Pharmacia micro quick prep mRNA extraction kit (Pharmacia, Piscataway NJ). cDNA will be prepared using mouse mammary leukemia virus (MMLV) and Superscript first strand cDNA synthesis kit (BRL) according to supplier's directions. Genomic DNA may also be analyzed by this technique. Both wild type and unknown DNA samples will be amplified using specific primers flanking the region of the mutation. These primers will contain the T7 or SP6 phage promoter consensus sequences respectively at their 5' ends followed by 15-18 bases of the target-specific sense or anti-sense flanking DNA sequences. Sense and anti-sense PCR products will be transcribed using T7 or SP6 polymerase. Complementary wild-type and patient transcripts will be hybridized. The resulting RNA duplexes will be digested with either RNase 1, RNase T1 and/or RNase A. Resulting products will be analyzed by separation on 2% agarose gel. Reagents and detailed protocols for this methodology are available in kit form (Ambion, Austin TX). This technique is capable of detecting both homozygous and heterozygous mutations.

Example 5

Description of a Putative Gene and Specific Mutations Linked to Absorptive Hypercalciuria

Example 2 describes genetic linkage between the clinical phenotype associated with absorptive hypercalciuria (AH) and the chromosome 1q24 locus. This example describes specific mutations of a putative gene located in this region and the relationship between this putative gene and AH. The frequency of mutations in this gene in patients with another disease, idiopathic osteoporosis, is also described, indicating a potential link between this gene and bone loss. The putative gene described here and the techniques described for elucidating the nature of this putative gene's role in AH are given as an

example. Should the putative gene described in this example eventually be shown not to be the AH gene, similar methodologies as described throughout this specification and proceeding examples will be used to identify the true AH gene.

Linkage studies and haplotype analysis localized a gene defect associated with AH to a 4.3 cM region of chromosome 1q between the markers D1S2681 centromere and D1S2815 (FIG. 2). As the highest lod scores were obtained between markers D1S196 and D1S431, this locus was initially chosen as the most likely location of the AH gene. Published sequence data from the Human Genome Project (www.sanger.ac.uk) identified the region of interest as chromosome 1 contig196. A large portion of this region is contained in 3 clones, namely dJ455J7 (GenBank Accession # AL031733) containing D1S196, and dJ313L4 (GenBank Accession # Z99943), and dJ295C6 (GenBank Accession # Z97876), both of which contain marker D1S431. Sequence data for these clones is available through GenBank (www.nih.gov). All clones contain numerous est's (expressed sequence tags) and therefore potential genes. No obvious candidate genes of known function are currently mapped to this region.

A 2567 base pair cDNA (GenBank Accession # AL035122) encoding a hypothetical protein of unknown function (GenBank Accession # CAA22684) has been mapped to the genomic region contained in dJ313L4 (GenBank Accession # Z99943) and dJ295C6 (GenBank Accession #Z97876). The nucleotide sequence of this cDNA is given as SEQ ID NO: 1, and the deduced amino acid sequence given as SEQ ID NO:2. BLAST alignment of the SEQ ID NO:1 and genomic sequence of dJ313L4 reveals a gene of at least 38,844 base pairs encompassing at least 16 exons (Bases 85943 to 124787 of dJ313L4, GenBank Accession # Z99943). The complete length of the gene is longer as the sequence for exon 1 is most likely incomplete and identification of the promoter sequences responsible for transcriptional control of this putative AH gene are yet to be determined.

Evidence for Association of the Putative Gene With the AH Defect

A probe was prepared by PCR amplification of genomic DNA using primers spanning the intron/exon boundaries for what is defined as exon 5 (Primer 1 CATCTAGGTTGCCTTACCCGAAGT, SEQ ID NO:5, primer 2 TGATTAGGAGCACAGCCTCAGTGC, SEQ ID NO:6) with the following amplification conditions: 10 minutes at 95°C, followed by 35 cycles of 94°C for 10 seconds, 63°C for 30 seconds and 72°C for 45 seconds and one final 10 minute incubation at 72°C. The probe was labeled with biotin and used to screen a human multiple tissue mRNA blot (Human RNA Master Blot, Catalog # 7770-1, Clonetech Laboratories, Palo Alto, CA), using conditions as described by the supplier. The results demonstrate high level expression of the putative AH gene in adult human colon, small intestine, kidney and liver. As the invariant features of AH are intestinal hyperabsorption of calcium and excessive urinary calcium excretion, intestine and kidney are target tissues for expression of a defective gene.

Sequence analysis of all exons and intron/exon boundaries of the putative gene in the probands from 2 of the kindred's revealed the same point mutation in what is defined as exon 5. The mutation was a C to A transversion mutation at position 823 of SEQ ID NO:1 and patients were heterozygous for the mutation. This is the expected finding as AH has a dominant mode of inheritance. This mutation occurs in the 5' nontranslated region of the of the putative open reading frame, 127 bases 5' to the intiator methionine. This particular mutation is not going to result in an amino acid change, but rather could affect translatability or stability/ half-life of the messenger RNA. This could lead to either an increased amount or decreased amount of the encoded protein, resulting in the AH phenotype.

The mutation destroys an Alu 1 restriction endonulease recognition site (AGCT to AGAT), thus providing a rapid RFLP screening method involving PCR amplification of the genomic DNA from a individual followed by Alu 1 restriction of the resulting PCR product. Cleavage of the PCR fragment by Alu 1 represents a wild type allele with

mutant alleles being resistant to Alu 1 cleavage. Preliminary analysis of the frequency of this mutation in normal, AH and idiopathic osteoporotic populations revealed evidence of a significantly higher occurrence of this mutation in both the AH and idiopathic osteoporotic populations (Table 7).

Table 7

Exon 5 Alu 1 RFLP, Occurrence in Control and Patient Populations.

	Control	AH	Idiopathic Osteoporotic
n	93	103	30
mutations	1	10*	4

*p = 0.02 (Fisher Exact Test)

Screening of all other exons will be undertaken using RNA mismatch cleavage analysis as described in the specification. For instance, analysis of AH patients using this method has identified a second mutation in what is defined as exon 4 of the putative gene. This second mutation is a T to C transition mutation at position 483 of SEQ ID NO:1 and patients were heterozygous for the mutation. Again, this is the expected finding as AH has a dominant mode of inheritance. This second mutation is not found in a number of normal control patients. In addition, analysis of 91 AH patients using this method identified 13 patients with two distinct mutations in what is defined as exon 2 of the putative gene. Sequence analysis of these patients will confirm the nature of the mutation. With each new mutation a population of normal individuals will be screened to confirm that it is not simply a polymorphism.

BLAST analysis of both the cDNA sequence, SEQ ID NO:1, and amino acid sequence, SEQ ID NO:2, revealed little homology with known sequences. However modeling studies of SEQ ID NO:2 predict a structure with 3 transmembrane domains. This is suggestive of a transport function. As the transport of calcium is enhanced in AH, this suggested structure would indicate a possible membrane ion transport protein. Due to the small size of the encoded protein (372 amino acids) it is likely that the gene may

encode a subunit of a larger multimeric protein. Recombinant expression of the putative gene in tissue culture cells followed by ion transport experiments will be conducted in an attempt to determine the function of the protein. Recombinant expression of the gene product will also enhance purification of the protein, generation of antibodies specific for the protein, the development of assay systems to allow biochemical studies of the proteins function, and development of screening assays for candidate compounds that modulate the proteins functions. Gene knockout in transgenic mice will be undertaken in an attempt to define the function of the gene in all organ systems and also the effect of lack of expression during embryonic development.

Example 6

Characterization of Other AH-related Modulators

To investigate the mechanisms involved in the transcriptional control of the AH gene, expression studies of the 5' flanking region of AH gene are performed. As many of the genes involved in the regulation of intestinal calcium absorption are regulated by vitamin D, this may include investigation of consensus vitamin D response elements in the AH locus.

A fragment of the AH-gene promoter region of interest is first inserted by standard cloning techniques into a reporter vector such as herpes simplex virus thymidine kinase promoter-containing vector pUTKAT3, which contains the chloramphenicol acetyl transferase (CAT) reporter gene (DeMay 1992). Transfection of the plasmid into GH4C1 cells is achieved using lipofection (GIBCO). 24 h before transfection the GH4C1 cells are fed with charcoal-stripped fetal calf serum (10% vol/vol in Dulbeccos modified Eagles medium (DMEM). In order to investigate the effect of vitamin D/vitamin D receptor-mediated transcriptional activity the cells are treated with 10^{-8} M 1,25 $(OH)_2$ D_3. Following a 24-48 h incubation period, CAT activity is assessed by thin layer chromatography (TLC). More specific analysis of interactions between various components of the vitamin D transcriptional modulation pathway is performed by transient co-transfection assays with the retinoid X receptor or other transcription factors.

Additionally, gel-retardation assays are performed to directly investigate any interaction between various regulatory proteins and the AH-gene promoter.

Identification of proteins interacting directly with the AH protein is best achieved by co-precipitation of a complex containing AH protein from human cells using a specific antibody directed against the AH-protein. This could be followed by purification of the modulator by standard protein chemistry techniques.

Example 7

Treatment of a Symptomatic or Asymptomatic Diagnosed Individual

Dietary and fluid regimens: Conservative dietary and fluid regimens should be incorporated into the daily routine of all patients with kidney stones. Fluid intake should typically be sufficient to produce 2-3 liters of urine /day. Additional fluid may be required in the summer or during exercise to compensate for fluid loss due to perspiration. Adequate hydration alone has been shown to decrease stone formation by as much as 60% (Hosking, 1983). Dietary modification should include restriction of sodium intake as high sodium intake results in increased urinary excretion of calcium and consequent increase in the saturation of stone-forming salts in urine. Intake of animal protein should be limited as this leads to an increase in bone resorption and a decrease in renal calcium resorption. Protein should be restricted to less than 8 ounces per day. There should also be a modest restriction of daily calcium intake in patients consuming a high calcium diet. However, patients with hypercalciuria with bone loss should not restrict calcium as this may lead to a negative calcium balance and potential worsening of bone problems. Certain food high in oxalate should also be restricted such as nuts, chocolate, brewed tea and green leafy vegetables. Daily consumption of citrus fruits or juices is beneficial due to their citrate, which directly binds calcium and also serves to increase urinary pH.

Therapeutic Measures: Therapeutic measures are directed towards reducing urinary calcium excretion and decreasing intestinal calcium bioavailability. Thiazides are

the first drug of choice: trichlormethiazide, 4 mg daily in a normal sized adult, hydrochlorothiazide, 25 mg twice daily or bendroflumethiazide, 2.5 mg twice daily. Hypokalemia and hypocitraturia are frequent complications of thiazide therapy and can be avoided by administration of potassium citrate, 15 to 20 mEq twice daily. Long term efficacy of thiazide therapy may be limited. It is recommended that in patients where hypocalciuric response is lost, a brief drug holiday of 6 months be instituted where sodium cellulose phosphate (SCP) therapy of 10-15 g/day in divided doses with meals is administered in place of thiazides. After completion of the SCP course thiazide responsiveness usually resumes. A new treatment, a slow-release, neutral potassium phosphate salt (UroPhos-K), has been shown to significantly reduce urinary calcium excretion with the added advantage of having sustained effectiveness. Efficacy is currently being assessed in a multi-institutional, double-blinded trial. In the future this may present a more beneficial treatment alternative. Detailed description of recommended medical therapy has been described by our group (Ruml *et al.*, 1997).

Medications: Certain medications are contraindicated; these include carbonic anhydrase inhibitors such as acetazolamide or methozolamide, ascorbic acid supplementation in patients with elevated urinary oxalate level, and trimethterene or trimethaterene-containing diuretics (Dyazide, Maxzide).

Treatment of asymptomatic diagnosed individuals: The emphasis of treatment is directed towards prevention of the first stone-forming episode. In all such individuals conservative dietary and fluid regiments should be followed as detailed in the previous section. Periodic urine analysis for evaluation of stone-forming salts and risk factors is recommended in addition. The frequency of these assessments should be determined by the individual's physician based on the evaluation of a baseline stone-risk profile. Potential worsening of risk factors can thus be assessed and, if necessary, therapeutic intervention can be introduced to prevent stone formation.

All of the methods and compositions disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the methods and/or compositions and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

REFERENCES

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

Abbondanzo *et al.*, *Breast Cancer Res. Treat.*, 16:182(#151), 1990.

Abrams, "Clinical studies of mineral metabolism in children using stable isotopes," *J. Ped. Gastroenter. Nutr.*, 19:151-163, 1994.

Allred *et al.*, *Breast Cancer Res. Treat.*, 16:182(#149), 1990.

Barkin, Wilson, Manuel, Bayley *et al.*, "Bone mineral content in idiopathic calcium nephrolithiasis," *Miner Electrol. Metab.*, 11:19-24, 1985.

Bataille, Achard, Fournier *et al.*, "Diet, vitamin D and vertebral mineral density in hypercalciuric calcium stone-formers," *Kidney Intern*, 39:1193-1205, 1991.

Bianchi, Vezzoli, Cusi *et al.*, "Abnormal red-cell calcium pump in patients with idiopathic hypercalciuria," *N. Engl. J. Med.*, 319:897-901, 1988.

Bordier *et al., Am. J. Med.* 63:398-409, 1977.

Breslau, Preminger, Adams, Otey, Pak "Use of ketoconazole to probe the pathogenic importance of 1,25-dihydroxyvitamin D in absorptive hypercalciuria," *J. Clin. Endocrinol Metab.*, 75:1446-1452, 1992.

Broadus, Insogna, Lang *et al.*, "Evidence for disordered control of 1,25-dihydroxyvitamin D production in absorptive hypercalciuria," *N. Engl. J. Med.*, 311:73-80, 1984.

Brown *et al.*, *Breast Cancer Res. Treat.*, 16:192(#191), 1990.

Brutlag *et al.*, *CABIOS,* 6:237-245, 1990.

Buck *et al., Br. J. Urol.* 53:485-491, 1981.

Campbell, in *Monoclonal Antibody Technology, Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 13, eds. Burden and Von Knippenberg, Amsterdam, Elseview, 75-83, 1984.

Chelly *et al., Proc. Natl. Acad. Sci.* 86:2617-2621, 1989.

Chou and Fasman, "Conformational Parameters for Amino Acids in Helical, β-Sheet, and Random Coil Regions Calculated from Proteins," *Biochemistry*, 13(2):211-222, 1974b.

Chou and Fasman, "Empirical Predictions of Protein Conformation," *Ann. Rev. Biochem.*, 47:251-276, 1978b.

Chou and Fasman, "Prediction of b-Turns," *Biophys. J.*, 26:367-384, 1979.

Chou and Fasman, "Prediction of Protein Conformation," *Biochemistry*, 13(2):222-245, 1974a.

Chou and Fasman, "Prediction of the Secondary Structure of Proteins from Their Amino Acid Sequence," *Adv. Enzymol. Relat. Areas Mol. Biol.*, 47:45-148, 1978a.

Coe, Parks, Moore, "Familial idiopathic hypercalciuria," *New Engl. J. Med.*, 300:337-340, 1979.

Collins, Frezal, Teague, Morton, "A metric map of humans:23,500 loci in 850 bands," *Proc. Natl. Acad. Sci. USA*, 93:1471-1475, 1996.

DeMay *et al., Proc. Natl. Acad. Sci.*, 89: 8097-8101, 1992.

Fetrow and Bryant, "New Programs for Potein Tertiary Structure Prediction," *Biotech.*, 11:479-483, 1993.

Frohman, In *PCR Protocols: A Guide to Methods and Applications*, 1990.

Fujita R. and Swaroop A., *BioTechniques* 18: 796-799, 1995.

Gefter *et al., Somatic Cell Genet.* 3:231-236, 1977.

Goding, in *Monoclonal Antibodies: Principles and Practice,* 2d ed., Orlando, Fla., Academic Press, 60-61, 65-66, 71-74, 1986.

Heller, Reza-Albarran, Breslau, Pak, "Sustained reduction in urinary calcium during long-term treatment with slow release neutral potassium phosphate in absorptive hypercalciuria," *J. Urol.*, 159:1451-1456, 1988.

Hess *et al., Am. J. Nephrol.* 13:18-26, 1993.

Horrigan and Westbrook, in *Gene Isolation and Mapping Protocols.* Human Press, ed. J. Boultwood, 123-125, 1997.

Hosking, *J. Urol.* 130:1115, 1983.

Inouye *et al., Nucleic Acids Res.*, 13:3101-3109, 1985.

Insogna, Broadus, Dreyer *et al.*, "Elevated production rate of 1,25-dihydroxyvitamin D in patients with absorptive hypercalciuria," *J. Clin. Endocrinol Metab.*, 61:490-495, 1985.

Jameson and Wolf, "The Antigenic Index: A Novel Algorithm for Predicting Antigenic Determinants," *Comput. Appl. Biosci.*, 4(1):181-186, 1988.

Johnson *et al.*, "Peptide Turn Mimetics," in *Biotechnology and Pharmacy*, Pezzuto *et al.*, eds., Chapman and Hall, New York, 1993.

Kaplan *et al., J. Clin. Invest.* 59:756-760, 1977.

Kasahara *et al.*, *Science*, 266:1373-1376, 1994.

Kohler and Milstein, *Eur. J. Immunol.*, 6:511-519, 1976.

Kohler and Milstein, *Nature*, 256:495-497, 1975.

Krieger *et al., Am. J. Physiol.* 271:C130-C135, 1996.

Krizman, in *Gene Isolation and Mapping Protocols*, Human Press, ed. J. Boultwood, 167-182, 1997.

Kruglyak, Daly, Reeve-Daly, Lander, "Parametric and non-parametric linkage analysis: a unified multilocus approach," *Am. J. Hum. Genet.*, 58:1347-1363, 1996.

Kyte and Doolittle, "A simple method for displaying the hydropathic character of a protein," *J. Mol. Biol.*, 157(1):105-132, 1982.

Lathrop, Lalouel, Julier, Ott, "Multilocus linkage analysis in humans: detection of linkage and estimation of recombination," *Am. J. Hum. Genet.*, 37: 482-498, 1985.

Ledbetter *et al., Genomics* 6: 475-481, 1990.

Levy, Adams-Huet, Pak, "Ambulatory evaluation of nephrolithiasis: an update of a 1980 protocol," Am. J. Med., 98:50-59, 1995.

Li, Tembe, Horowitz, Bushinsky, Favus, "Increased intestinal vitamin D receptor in genetic hypercalciuric rats: a cause of intestinal calcium hyperabsorption," *J. Clin. Invest.*, 91:661-667, 1993.

Lloyd, Pearce, Fisher *et al.*, "A common molecular basis for three inherited kidney stone diseases," *Nature*, 379:445-449, 1996.

Malluche *et al., J. Clin Endocrin. Metab.* 50:654-658, 1980.

Miller, *Curr. Top. Microbiol. Immunol.*, 158:1, 1992.

Nakamura *et al.*, *In*: *Enzyme Immunoassays: Heterogeneous and Homogeneous Systems*, Chapter 27.

Neitzel, "A routine method for the establishment of permanent growing lymphoblastoid cell lines," *Hum. Genet.*, 73:320-326, 1989.

Pacifici *et al.*, *J. Clin. Endo. Metab.*, 71:138, 1990.

Pak and Galosy, "Fasting urinary calcium and adenosine 3', 5,-monophosphate: a discriminant analysis for the identification of renal and absorptive hypercalciurias," *J. Clin. Endocrinol. Metab.*, 48: 260-265, 1979.

Pak, Britton, Peterson *et al.*, "Ambulatory evaluation of nephrolithiasis: classification clinical presentation, and diagnostic criteria," *Am. J. Med.*, 69:19-30, 1980.

Pak, Kaplan, Bone *et al.*, "A simple test for the diagnosis of absorptive, resorptive and renal hypercalciurias," *N. Engl. J. Med.*, 292:497-500, 1975.

Pak, Mc Guire, Peterson *et al.*, "Familial absorptive hypercalciuria in a large kindred," *J. Urology*, 126:717-721, 1981.

Pak, Ohata, Lawrence, Snyder, "The hypercalciurias: causes, parathyroid functions and diagnostic criteria," *J. Clin. Invest.*, 54:387-400, 1974.

Pietchmann, Breslau, Pak, "Reduced vertebral bone density in hypercalciuric nephrolithiasis," *J. Bone Min. Res.*, 7:1383-1388, 1992.

Preminger, Peterson, Pak, "Differentiation of unclassified hypercalciuria utilizing a sodium cellulose phosphate trial," *In Urolithiasis,* Walker, Sutton, Cameron, Pak, Robertson (eds.), Plenum Press New York and London, 325-328, 1989.

Reed, Heller, Lemke *et al.*, "Linkage analysis in absorptive hypercalciuria: lack of linkage to the vitamin D receptor or $1,25(OH)_2D_3$ hydroxylase loci," *In Urolithiasis*, Pak, Resnick, Preminger (eds), Millet Press, Dallas, TX, 540-542, 1996.

Resnick, Prigden, Goodman, "Genetic predisposition to formation of calcium oxalate renal calculi," *N. Engl. J. Med.*, 278:1313-1318, 1968.

Ruml *et al.*, *Urol. Clinics of N. Am.* 24:117-132, 1997.

Sambrook, Fritsch, Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989.

Steiniche *et al.*, *APMIS* 97:309-316, 1989.

Tenenhouse, "Cellular and molecular mechanisms of renal phosphate transport," *J. Bone Min. Res.*, 12:159-164, 1997.

Tieder, Modat, Samuel, *et al.*, "Heriditary hypophosphatemic rickets with hypercalciuria," *N. Eng. J. Med.*, 312:611-617, 1985.

Weinberger *et al.*, *Science,* 228:740-742, 1985.

Weisinger *et al. Kidney Int.* 49:244-250, 1996.

Wolf *et al.*, "An Integrated Family of Amino Acid Sequence Analysis Programs," *Comput. Appl. Biosci.*, 4(1):187-191, 1988.

Wong *et al.*, "Appearance of β-lactamase activity in animal cells upon liposome mediated gene transfer," *Gene*, 10:87-94, 1980.

Zerwekh, Hughes, Reed *et al.*, "Evidence for normal vitamin D receptor messenger ribonucleic acid and genotype in absorptive hypercalciuria," *J. Clin. Endocrinol. Metab.*, 80:2960-2965, 1995.

Zerwekh, Yu, Breslau, Manologas, Pak, "Vitamin D receptor quantitation in human mononuclear cells in health and disease," *Mol. Cell Endocrinol.*, 96:1-6, 1993.

TAS A17